XX
PD 11-OCT-2001.
XX
PF 03-APR-2001; 2001WO-US010908.
XX
PR 03-APR-2000; 2000US-0194504P.
PR 28-NOV-2000; 2000US-0253647P.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
PA (PFIZ) PFIZER INC.
XX
PI Durham KL, Friedman DL, Herath HMAC, Kimmel LH, Parekh RB;
PI Potter DM, Rohlff C, Silber BM, Stiger TR, Sunderland PT;
PI Townsend RR, White F, Williams SA;
XX
DR WPI; 2001-639384/73.
XX
PT Screening for Alzheimer's disease in a mammal, by making two-dimensional
PT array of a feature whose relative abundance correlates with disease, and
PT comparing with abundance of the feature in samples of healthy persons.
XX
PS Example; Page 30; 162pp; English.
XX
CC The invention relates to methods for the screening, diagnosis and
CC prognosis of Alzheimer's disease. The methods involve the detection of
CC Alzheimer's Disease-Associated Features (AFs) and Alzheimer's Disease-
CC Associated Protein Isoforms (APIs) in cerebrospinal fluid, serum or
CC plasma. The abundance of the AFs and APIs is then normalised to an
CC Expression Reference Protein Isoform (ERPI) in order to determine whether
CC a patient is suffering from, or has a predisposition to, Alzheimer's
CC Disease. The relative abundance of the AFs and APIs correlates with the
CC severity of Alzheimer's Disease. The present sequence is a peptide
CC produced from an API by proteolysis
XX
SQ Sequence 7 AA;

Query Match 60.6%; Score 20; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 FTLKISR 7
              || : ||
Db          1 FTFEYSR 7
```

Search completed: June 1, 2004, 11:16:18
Job time : 47.6667 secs

KW CSF; antidepressant; antimanic; nootropic; tranquiliser; neuroleptic;
KW attention deficient disorder; schizoaffective disorder;
KW unipolar affective disorder.
XX
OS Homo sapiens.
XX
PN WO200163294-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000791.
XX
PR 24-FEB-2000; 2000GB-00004412.
PR 08-DEC-2000; 2000GB-00030050.
PR 12-DEC-2000; 2000US-0254830P.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-582081/65.
XX
PT Preparation for diagnosing or treating bipolar affected disorder (BAD) or
PT unipolar depression, or for screening for modulators, comprises a BAD-
PT associated protein isoform.
XX
PS Claim 8; Page 34; 163pp; English.
XX
CC The invention relates to a preparation comprising an isolated Bipolar
CC Affected Disorder (BAD)-Associated Protein Isoform (DPIs). The DPI's are
CC used to screen, diagnose or prognose of BAD or unipolar depression,
CC determine the stage or severity of BAD or unipolar depression, identify a
CC subject at risk of developing BAD or unipolar depression, or monitor the
CC effect of therapy in a subject. They are also used to screen for or
CC identify agents that interact with a DPI. These agents, antibodies
CC against the DPIs, and nucleic acids encoding the DPIs are used to treat
CC or prevent BAD or unipolar depression. Diseases that can be treated are
CC attention deficient disorder, a schizoaffective disorder, a bipolar or a
CC unipolar affective disorder. The DPIs are used in proteomics. The
CC proteomic approach of using DPIs for screening, diagnosis or prognosis of
CC BAD or unipolar depression overcomes the problems of using gene
CC expression analysis, such as not being able to obtain central nervous
CC system (CNS) tissue from a living patient under normal circumstances. The
CC present sequence is a DIP decreased in the CSF (cerebro-spinal fluid) of
CC subjects having BAD
XX
SQ Sequence 7 AA;

Query Match 60.6%; Score 20; DB 4; Length 7;
Best Local Similarity 42.9%; Pred. No. 1.4e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLKISR 7
:| ::||
Db 1 YTFELSR 7

RESULT 14
AAU15313
ID AAU15313 standard; peptide; 7 AA.
XX
AC AAU15313;
XX
DT 24-OCT-2001 (first entry)
XX
DE Schizophrenia-associated isoform peptide #198.
XX
KW Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder;
KW neurological disorder; neuropathy.
XX
OS Homo sapiens.
XX
PN WO200163293-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000783.
XX
PR 24-FEB-2000; 2000GB-00004415.
PR 28-DEC-2000; 2000US-00750395.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-502868/55.
XX
PT Diagnosing and monitoring Schizophrenia by detecting the presence of
PT Schizophrenia Associated Features and Schizophrenia Associated Protein
PT Isoforms in samples of cerebrospinal fluid.
XX
PS Claim 6; Page 32; 160pp; English.
XX
CC The invention relates to methods and compositions for screening,
CC diagnosis and prognosis of Schizophrenia. The method involves detecting
CC the presence of Schizophrenia (SCH) Associated Features (SFs) and SCH
CC Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis,
CC immunoassay or hybridisation assay, for diagnosing and monitoring SCH,
CC studying the effectiveness of treatments and for identifying potential
CC therapeutic agents. The method is used for (1) screening or diagnosis of
CC SCH and the relative abundance of at least 1 chosen feature correlates
CC with the presence or absence of SCH; and (2) monitoring the effect of
CC therapy administered to a subject with SCH and the relative abundance of
CC at least 1 chosen feature which correlates with the severity of SCH. The
CC expression and activity of the SFs, SPIs and related molecules (e.g.
CC secondary messengers) are studied to diagnose SCH, monitor the progress
CC of the disorder and the effectiveness of treatment and as targets to
CC identify and produce potential therapeutic agents for the treatment of
CC SCH. The paucity of detectable neuralgic defects distinguishes
CC neuropsychiatric disorders such as SCH from neurological disorders, where
CC manifestations of anatomical and biochemical changes have been identified
CC in many cases. Consequently the identification and characterisation of
CC cellular and/or molecular causative defects and neuropathies are
CC necessary for improved treatment of neuropsychiatric disorders. AAU15114-
CC AAU15762 represent the amino acid sequences of schizophrenia-associated
CC isoforms used in the method of the invention
XX
SQ Sequence 7 AA;

Query Match 60.6%; Score 20; DB 4; Length 7;
Best Local Similarity 42.9%; Pred. No. 1.4e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLKISR 7
:| ::||
Db 1 YTFELSR 7

RESULT 15
ABB52190
ID ABB52190 standard; peptide; 7 AA.
XX
AC ABB52190;
XX
DT 08-FEB-2002 (first entry)
XX
DE Human API-146 tryptic digest peptide #1.
XX
KW Human; neuroprotective; nootropic; gene therapy; vaccine;
KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;
KW Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;
KW Expression Reference Protein Isoform; ERPI; proteolysis.
XX
OS Homo sapiens.
XX
PN WO200175454-A2.

```
CC   treatment of VD and for gene therapy
XX
SQ   Sequence 7 AA;

  Query Match             60.6%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1.4e+06;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              :| ::||
Db          1 YTFELSR 7


RESULT 11
AAU28602
ID   AAU28602 standard; peptide; 7 AA.
XX
AC   AAU28602;
XX
DT   03-JAN-2002  (first entry)
XX
DE   DPI tryptic digest peptide #199.
XX
KW   Human; depression associated protein isoform; tryptic digest peptide;
KW   DPI; cerebrospinal fluid; CSF; BAD; bipolar affective disorder;
KW   neuropsychiatric disorder; bipolar mood disorder; neuroleptic;
KW   maniac-depressive illness; schizoaffective disorder.
XX
OS   Homo sapiens.
XX
PN   WO200162787-A1.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000786.
XX
PR   24-FEB-2000; 2000GB-00004412.
PR   08-DEC-2000; 2000GB-00030050.
PR   12-DEC-2000; 2000US-0254830P.
XX
PA   (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI   Herath HMAC,  Parekh RB,  Rohlff C,  Terrett JA,  Tyson KL;
XX
DR   WPI; 2001-570626/64.
XX
PT   Novel nucleic acid encoding a protein associated with bipolar affective
PT   disorder, which is used for diagnosis, prophylaxis and therapy of
PT   neuropsychiatric disorders, such as bipolar affective disorder.
XX
PS   Disclosure; Page 34; 153pp; English.
XX
CC   The present invention relates to the identification of depression
CC   associated protein isoforms (DPIs), particularly the tryptic digest
CC   peptides of these proteins. Some of the DPIs (AAU28404-AAU28625)
CC   described are decreased in the cerebrospinal fluid (CSF) of BAD (bipolar
CC   affective disorder) subjects, whilst other DPIs (AAU28626-AAU28887) are
CC   increased in BAD subjects. Also described are peptide sequences
CC   identified from DPI-45 and DPI-213 and the nucleic acid sequence they are
CC   encoded by. The sequences of the invention are useful for clinical
CC   screening, diagnosis, prognosis, therapy and prophylaxis of
CC   neuropsychiatric disorders e.g. BAD (also known as bipolar mood disorder,
CC   BP), maniac-depressive illnesses, attention deficit disorders,
CC   schizoaffective disorders, and unipolar affective disorders. The present
CC   sequence represents one of the DPI tryptic digest peptides of the present
CC   invention
XX
SQ   Sequence 7 AA;

  Query Match             60.6%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1.4e+06;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              :| ::||
Db          1 YTFELSR 7


RESULT 12
AAU24969
ID   AAU24969 standard; peptide; 7 AA.
XX
AC   AAU24969;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Schizophrenia-Associated Protein Isoform (SPI) peptide #198.
XX
KW   Schizophrenia-associated protein isoform; SPI; SPI-206; SPI-238; SPI-240;
KW   neuroleptic; gene therapy; cerebrospinal fluid; serum; plasma.
XX
OS   Homo sapiens.
XX
PN   WO200162785-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000792.
XX
PR   24-FEB-2000; 2000GB-00004415.
PR   28-DEC-2000; 2000US-00750395.
XX
PA   (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI   Herath HMAC,  Parekh RB,  Rohlff C,  Terrett JA,  Tyson KL;
XX
DR   WPI; 2001-570624/64.
XX
PT   New schizophrenia associated protein isoforms and encoding nucleic acid
PT   molecules, useful for treatment, diagnosis and prognosis of schizophrenia
PT   and screening for potential drugs for treatment and new drug targets.
XX
PS   Disclosure; Page 32; 148pp; English.
XX
CC   The sequence represents a schizophrenia-associated protein isoform (SPI).
CC   These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable
CC   in cerebrospinal fluid, serum or plasma and are useful markers of
CC   schizophrenia. The sequences can be used for treatment and diagnosis of
CC   schizophrenia, screening, prognosis, monitoring the results of therapy,
CC   identifying patients most likely to respond to a particular therapy and
CC   identification of new targets for drug treatment. SPI DNA is useful as a
CC   nucleic acid probe to detect the presence of nucleic acids or SPIs
XX
SQ   Sequence 7 AA;

  Query Match             60.6%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1.4e+06;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              :| ::||
Db          1 YTFELSR 7


RESULT 13
AAU26249
ID   AAU26249 standard; peptide; 7 AA.
XX
AC   AAU26249;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Depression-Associated Protein isoform DPI-208.
XX
KW   Human; Bipolar Affective Disorder; BAD; Depression-Associated feature;
KW   DF; Depression-Associated protein isoform; DPI; Cerebro-spinal fluid;
```

PT comprises analyzing body fluid by 2-dimensional electrophoresis for
PT features correlated with VD.
XX
PS Claim 6; Page 31; 151pp; English.
XX
CC The invention relates to screening, diagnosis or prognosis of Vascular
CC Dementia (VD) in a subject comprising analysing body fluid from the
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC features containing at least one chosen feature whose relative abundance
CC correlates with the presence, absence, stage or severity of VD or
CC predicts the onset or course of VD, especially detecting in a sample of
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC specification. Detecting VD-associated features and VPI is useful for the
CC screening, diagnosis or prognosis of VD, for determining the stage or
CC severity of VD, for identifying a subject at risk of VD or for monitoring
CC the effect of therapy administered to a subject having VD. Nucleic acids
CC encoding a VPI or inhibiting the function of a VPI are useful for the
CC treatment of VD and for gene therapy
XX
SQ Sequence 7 AA;

Query Match 60.6%; Score 20; DB 4; Length 7;
Best Local Similarity 42.9%; Pred. No. 1.4e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLKISR 7
:| ::||
Db 1 YTFELSR 7

RESULT 9
ABB56283
ID ABB56283 standard; peptide; 7 AA.
XX
AC ABB56283;
XX
DT 15-FEB-2002 (first entry)
XX
DE Vascular dementia-associated protein isoform (VPI) 483.
XX
KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW diagnosis; prognosis; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200169261-A2.
XX
PD 20-SEP-2001.
XX
PF 14-MAR-2001; 2001WO-GB001106.
XX
PR 15-MAR-2000; 2000GB-00006285.
PR 24-NOV-2000; 2000GB-00028734.
PR 28-NOV-2000; 2000US-00724391.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-557937/62.
XX
PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT determining stage of VD and monitoring the effect of VD therapy,
PT comprises analyzing body fluid by 2-dimensional electrophoresis for
PT features correlated with VD.
XX
PS Claim 6; Page 40; 151pp; English.
XX
CC The invention relates to screening, diagnosis or prognosis of Vascular
CC Dementia (VD) in a subject comprising analysing body fluid from the
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC features containing at least one chosen feature whose relative abundance
CC correlates with the presence, absence, stage or severity of VD or
CC predicts the onset or course of VD, especially detecting in a sample of
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC specification. Detecting VD-associated features and VPI is useful for the
CC screening, diagnosis or prognosis of VD, for determining the stage or
CC severity of VD, for identifying a subject at risk of VD or for monitoring
CC the effect of therapy administered to a subject having VD. Nucleic acids
CC encoding a VPI or inhibiting the function of a VPI are useful for the
CC treatment of VD and for gene therapy
XX
SQ Sequence 7 AA;

Query Match 60.6%; Score 20; DB 4; Length 7;
Best Local Similarity 42.9%; Pred. No. 1.4e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLKISR 7
:| ::||
Db 1 YTFELSR 7

RESULT 10
ABB55981
ID ABB55981 standard; peptide; 7 AA.
XX
AC ABB55981;
XX
DT 15-FEB-2002 (first entry)
XX
DE Vascular dementia-associated protein isoform (VPI) 181.
XX
KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW diagnosis; prognosis; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200169261-A2.
XX
PD 20-SEP-2001.
XX
PF 14-MAR-2001; 2001WO-GB001106.
XX
PR 15-MAR-2000; 2000GB-00006285.
PR 24-NOV-2000; 2000GB-00028734.
PR 28-NOV-2000; 2000US-00724391.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-557937/62.
XX
PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT determining stage of VD and monitoring the effect of VD therapy,
PT comprises analyzing body fluid by 2-dimensional electrophoresis for
PT features correlated with VD.
XX
PS Claim 6; Page 33; 151pp; English.
XX
CC The invention relates to screening, diagnosis or prognosis of Vascular
CC Dementia (VD) in a subject comprising analysing body fluid from the
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC features containing at least one chosen feature whose relative abundance
CC correlates with the presence, absence, stage or severity of VD or
CC predicts the onset or course of VD, especially detecting in a sample of
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC specification. Detecting VD-associated features and VPI is useful for the
CC screening, diagnosis or prognosis of VD, for determining the stage or
CC severity of VD, for identifying a subject at risk of VD or for monitoring
CC the effect of therapy administered to a subject having VD. Nucleic acids
CC encoding a VPI or inhibiting the function of a VPI are useful for the

XX
PF 31-MAR-1995; 95WO-US004075.
XX
PR 31-MAR-1994; 94US-00222619.
XX
PA (AMGE-) AMGEN INC.
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Lichenstein HS, Lyons DE, Wurfel MM, Wright SD;
XX
DR WPI; 1995-358634/46.
XX
PT Human afamin or a variant and poly:nucleotide(s) encoding it - a human
PT serum protein with activities in common with other members of this
PT family.
XX
PS Example 3; Page 45; 97pp; English.
XX
CC Peptides AAR81847-54 are tryptic peptide fragments from human afamin
CC (AAR81845) novel member of the human serum protein family. The fragments
CC were used to design primers and probes (AAT00786-98) for the cloning of
CC the afamin gene (AAT00785) from human liver cDNA. Afamin is thought to
CC have similar properties to human albumin, alpha-foetoprotein and vitamin
CC D binding protein due to homology with these proteins. The gene encodes a
CC mature protein of 66576 daltons without post-translational processing
CC (ca. 87000 daltons with post-translational processing). The protein was
CC isolated from human plasma by a conventional chromatographic methods. The
CC protein can be used to ameliorate ischaemia-reperfusion injury,
CC rheumatoid arthritis, ARDS, cardiopulmonary bypass, sepsis, toxic plasma
CC substances released after inflammation, etc
XX
SQ Sequence 7 AA;

Query Match 60.6%; Score 20; DB 2; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FTLKISR 7
|| : ||
Db 1 FTFEYSR 7

RESULT 7
AAY41889
ID AAY41889 standard; peptide; 7 AA.
XX
AC AAY41889;
XX
DT 09-DEC-1999 (first entry)
XX
DE Rheumatoid arthritis diagnostic protein isoform peptide #40.
XX
KW Human; rheumatoid arthritis; RA; diagnosis; RPI; RADF; detection;
KW rheumatoid arthritis diagnostic feature; ERPI; synovial fluid;
KW rheumatoid arthritis diagnostic protein isoform; screening;
KW expression reference protein isoform; prognosis.
XX
OS Homo sapiens.
XX
PN WO9947925-A2.
XX
PD 23-SEP-1999.
XX
PF 15-MAR-1999; 99WO-GB000763.
XX
PR 13-MAR-1998; 98GB-00005477.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Parekh RB, Patel TP, Townsend RR;
XX
DR WPI; 1999-571871/48.
XX
PT Diagnosis of human rheumatoid arthritis by two-dimensional
PT electrophoresis.
XX
PS Disclosure; Page 18; 157pp; English.
XX
CC A method has been developed for the diagnosis of human rheumatoid
CC arthritis (RA) using two-dimensional electrophoresis to generate a two-
CC dimensional array of features. The method can be used for screening,
CC diagnosis and prognosis of RA in a subject or for monitoring the effect
CC of an anti-RA drug or therapy administered to a subject. The method
CC comprises: (a) analysing a sample of serum or plasma and optionally
CC synovial fluid by two-dimensional electrophoresis, to generate a two-
CC dimensional array of features; (b) identifying at least one chosen
CC feature whose relative abundance correlates with the presence or absence
CC of RA; and (c) comparing the abundance of each chosen feature in the
CC sample with the abundance of that chosen feature in serum or plasma from
CC one or more persons without RA, where the relative abundance of the
CC chosen feature or features in the sample indicates the presence or
CC absence of RA in the subject. The method can also be used in clinical
CC studies for testing drugs for therapy of RA, for purification of RA-
CC diagnostic protein isoforms (RPIs), and for production of antibodies to
CC RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify
CC compounds that promote or inhibit their activity, which are then used as
CC RA drugs. Nucleic acid encoding RADFs can be used in gene therapy
CC protocols. AAY41844 to AAY42100 represent RPI peptides, AAY42101 to
CC AAY42103 represent expression reference protein isoform peptides and
CC AAZ25066 to AAZ25068 represent degenerate probes for RPIs, which are all
CC used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 60.6%; Score 20; DB 2; Length 7;
Best Local Similarity 42.9%; Pred. No. 1.4e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLKISR 7
:| ::||
Db 1 YTFELSR 7

RESULT 8
ABB55870
ID ABB55870 standard; peptide; 7 AA.
XX
AC ABB55870;
XX
DT 15-FEB-2002 (first entry)
XX
DE Vascular dementia-associated protein isoform (VPI) 70.
XX
KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW diagnosis; prognosis; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200169261-A2.
XX
PD 20-SEP-2001.
XX
PF 14-MAR-2001; 2001WO-GB001106.
XX
PR 15-MAR-2000; 2000GB-00006285.
PR 24-NOV-2000; 2000GB-00028734.
PR 28-NOV-2000; 2000US-00724391.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-557937/62.
XX
PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT determining stage of VD and monitoring the effect of VD therapy,

AAB30076
ID AAB30076 standard; peptide; 7 AA.
XX
AC AAB30076;
XX
DT 09-FEB-2001 (first entry)
XX
DE Scaffold protein SCA S4 peptide SEQ ID NO: 137.
XX
KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.
XX
OS Synthetic.
XX
PN WO200060070-A1.
XX
PD 12-OCT-2000.
XX
PF 01-APR-1999; 99WO-EP002283.
XX
PR 01-APR-1999; 99WO-EP002283.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 2000-665002/64.
XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding fragments.
XX
PS Disclosure; Page 15; 68pp; English.
XX
CC The present invention is concerned with producing scaffold proteins based
CC upon the human CTLA-4 SCA domain. These scaffold proteins can be used as
CC a scaffold to bind antigen- or receptor-binding fragments. These can be
CC used in the treatment of diseases such as cancer, atherosclerosis,
CC thrombosis, osteoporosis, rheumatoid arthritis and diabetic retinopathy.
CC Sequences AAB29930-B29939 were used in the production of the proteins of
CC the invention
XX
SQ Sequence 7 AA;

Query Match 66.7%; Score 22; DB 3; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTLKIS 6
              ||| ||
Db          2 FTLTIS 7
```

RESULT 5
AAY42013
ID AAY42013 standard; peptide; 7 AA.
XX
AC AAY42013;
XX
DT 09-DEC-1999 (first entry)
XX
DE Rheumatoid arthritis diagnostic protein isoform peptide #164.
XX
KW Human; rheumatoid arthritis; RA; diagnosis; RPI; RADF; detection;
KW rheumatoid arthritis diagnostic feature; ERPI; synovial fluid;
KW rheumatoid arthritis diagnostic protein isoform; screening;
KW expression reference protein isoform; prognosis.
XX
OS Homo sapiens.
XX
PN WO9947925-A2.
XX
PD 23-SEP-1999.
XX
PF 15-MAR-1999; 99WO-GB000763.
XX
PR 13-MAR-1998; 98GB-00005477.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Parekh RB, Patel TP, Townsend RR;
XX
DR WPI; 1999-571871/48.
XX
PT Diagnosis of human rheumatoid arthritis by two-dimensional
PT electrophoresis.
XX
PS Disclosure; Page 21; 157pp; English.
XX
CC A method has been developed for the diagnosis of human rheumatoid
CC arthritis (RA) using two-dimensional electrophoresis to generate a two-
CC dimensional array of features. The method can be used for screening,
CC diagnosis and prognosis of RA in a subject or for monitoring the effect
CC of an anti-RA drug or therapy administered to a subject. The method
CC comprises: (a) analysing a sample of serum or plasma and optionally
CC synovial fluid by two-dimensional electrophoresis, to generate a two-
CC dimensional array of features; (b) identifying at least one chosen
CC feature whose relative abundance correlates with the presence or absence
CC of RA; and (c) comparing the abundance of each chosen feature in the
CC sample with the abundance of that chosen feature in serum or plasma from
CC one or more persons without RA, where the relative abundance of the
CC chosen feature or features in the sample indicates the presence or
CC absence of RA in the subject. The method can also be used in clinical
CC studies for testing drugs for therapy of RA, for purification of RA-
CC diagnostic protein isoforms (RPIs), and for production of antibodies to
CC RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify
CC compounds that promote or inhibit their activity, which are then used as
CC RA drugs. Nucleic acid encoding RADFs can be used in gene therapy
CC protocols. AAY41844 to AAY42100 represent RPI peptides, AAY42101 to
CC AAY42103 represent expression reference protein isoform peptides and
CC AAZ25066 to AAZ25068 represent degenerate probes for RPIs, which are all
CC used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 63.6%; Score 21; DB 2; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 TLKISR 7
              || |||
Db          2 TLMISR 7
```

RESULT 6
AAR81848
ID AAR81848 standard; peptide; 7 AA.
XX
AC AAR81848;
XX
DT 16-MAY-1996 (first entry)
XX
DE Human afamin tryptic fragment FX20.
XX
KW Human; afamin; serum protein family; albumin; alpha-foetoprotein; plasma;
KW vitamin D binding protein; homology; post-translational processing;
KW chromatography; Primer; PCR; amplification; probe; rheumatoid arthritis;
KW ischaemia-reperfusion injury; ARDS; cardiopulmonary bypass; sepsis;
KW toxic plasma substance; inflammation.
XX
OS Homo sapiens.
XX
PN WO9527059-A1.
XX
PD 12-OCT-1995.

CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines
XX
SQ Sequence 7 AA;

Query Match 69.7%; Score 23; DB 2; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLKIS 6
||| ||
Db 2 FTLSIS 7

RESULT 2
AAB30074
ID AAB30074 standard; peptide; 7 AA.
XX
AC AAB30074;
XX
DT 09-FEB-2001 (first entry)
XX
DE Scaffold protein SCA S4 peptide SEQ ID NO: 135.
XX
KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.
XX
OS Synthetic.
XX
PN WO200060070-A1.
XX
PD 12-OCT-2000.
XX
PF 01-APR-1999; 99WO-EP002283.
XX
PR 01-APR-1999; 99WO-EP002283.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 2000-665002/64.
XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding fragments.
XX
PS Disclosure; Page 15; 68pp; English.
XX
CC The present invention is concerned with producing scaffold proteins based
CC upon the human CTLA-4 SCA domain. These scaffold proteins can be used as
CC a scaffold to bind antigen- or receptor-binding fragments. These can be
CC used in the treatment of diseases such as cancer, atherosclerosis,
CC thrombosis, osteoporosis, rheumatoid arthritis and diabetic retinopathy.
CC Sequences AAB29930-B29939 were used in the production of the proteins of
CC the invention
XX
SQ Sequence 7 AA;

Query Match 69.7%; Score 23; DB 3; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLKIS 6
||| ||
Db 2 FTLSIS 7

RESULT 3
AAY40738
ID AAY40738 standard; peptide; 7 AA.
XX
AC AAY40738;
XX
DT 01-DEC-1999 (first entry)
XX
DE S4 derivative #12, beta strand of scaffold protein structure.
XX
KW Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.
XX
OS Synthetic.
XX
PN EP947582-A1.
XX
PD 06-OCT-1999.
XX
PF 31-MAR-1998; 98EP-00870065.
XX
PR 31-MAR-1998; 98EP-00870065.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 1999-542958/46.
XX
PT New scaffold protein, useful for stabilizing antigens used as vaccines.
XX
PS Disclosure; Page 6; 105pp; English.
XX
CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides (AAY40601-
CC Y40609) together form a single-chain scaffold protein which contains at
CC least 1 disulfide bond, contains less than 10% alpha helix and contains
CC at least 6 beta-strands. The scaffold protein is constructed of beta
CC strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines
XX
SQ Sequence 7 AA;

Query Match 66.7%; Score 22; DB 2; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLKIS 6
||| ||
Db 2 FTLTIS 7

RESULT 4

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:40:31 ; Search time 45.6667 Seconds
(without alignments)
43.310 Million cell updates/sec

Title: US-09-712-819C-6
Perfect score: 33
Sequence: 1 FTLKISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 92273

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	23	69.7	7	2	AAY40736	Aay40736 S4 deriva
2	23	69.7	7	3	AAB30074	Aab30074 Scaffold
3	22	66.7	7	2	AAY40738	Aay40738 S4 deriva
4	22	66.7	7	3	AAB30076	Aab30076 Scaffold
5	21	63.6	7	2	AAY42013	Aay42013 Rheumatoi
6	20	60.6	7	2	AAR81848	Aar81848 Human afa
7	20	60.6	7	2	AAY41889	Aay41889 Rheumatoi
8	20	60.6	7	4	ABB55870	Abb55870 Vascular
9	20	60.6	7	4	ABB56283	Abb56283 Vascular
10	20	60.6	7	4	ABB55981	Abb55981 Vascular
11	20	60.6	7	4	AAU28602	Aau28602 DPI trypt
12	20	60.6	7	4	AAU24969	Aau24969 Schizophr
13	20	60.6	7	4	AAU26249	Aau26249 Depressio
14	20	60.6	7	4	AAU15313	Aau15313 Schizophr
15	20	60.6	7	4	ABB52190	Abb52190 Human API
16	20	60.6	7	4	ABB52355	Abb52355 Human API
17	20	60.6	7	5	ABG78901	Abg78901 Multiple
18	20	60.6	7	5	ABG78730	Abg78730 Multiple
19	20	60.6	7	6	ABP58010	Abp58010 Prostate
20	20	60.6	7	6	ABP57255	Abp57255 Breast ca
21	20	60.6	7	6	ABP57203	Abp57203 Breast ca
22	20	60.6	7	6	ABR59010	Abr59010 Alzheimer
23	20	60.6	7	6	ABR59042	Abr59042 Alzheimer
24	19	57.6	7	2	AAY40737	Aay40737 S4 deriva
25	19	57.6	7	3	AAB30075	Aab30075 Scaffold
26	19	57.6	7	7	ADB79629	Adb79629 Parapoxvi
27	18	54.5	7	5	ABP66518	Abp66518 Human RSV
28	18	54.5	7	5	ABR40481	Abr40481 Human sec
29	18	54.5	7	5	ABB81854	Abb81854 Staphyloc
30	18	54.5	7	6	ABU69381	Abu69381 Respirato
31	18	54.5	7	7	ADE35876	Ade35876 SYNAGIS a
32	18	54.5	7	7	ADE77927	Ade77927 Synthetic
33	17	51.5	6	2	AAW84431	Aaw84431 HIV-1 nuc
34	17	51.5	6	5	AAE28110	Aae28110 Human imm
35	17	51.5	6	6	ABP95994	Abp95994 Microtetr
36	17	51.5	7	2	AAR07656	Aar07656 Ribonucle
37	17	51.5	7	2	AAR97970	Aar97970 Antigenic
38	17	51.5	7	2	AAW69269	Aaw69269 Haemagglu
39	17	51.5	7	5	ABP66489	Abp66489 Human RSV
40	17	51.5	7	6	ABU69352	Abu69352 Respirato
41	17	51.5	7	7	ADE35847	Ade35847 SYNAGIS a
42	17	51.5	7	7	ADE77919	Ade77919 Synthetic
43	17	51.5	7	7	ADE78059	Ade78059 Synthetic
44	16	48.5	5	2	AAR75584	Aar75584 gp120 bin
45	16	48.5	6	1	AAP93345	Aap93345 Portion o

ALIGNMENTS

RESULT 1
AAY40736
ID AAY40736 standard; peptide; 7 AA.
XX
AC AAY40736;
XX
DT 01-DEC-1999 (first entry)
XX
DE S4 derivative #10, beta strand of scaffold protein structure.
XX
KW Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.
XX
OS Synthetic.
XX
PN EP947582-A1.
XX
PD 06-OCT-1999.
XX
PF 31-MAR-1998; 98EP-00870065.
XX
PR 31-MAR-1998; 98EP-00870065.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 1999-542958/46.
XX
PT New scaffold protein, useful for stabilizing antigens used as vaccines.
XX
PS Disclosure; Page 6; 105pp; English.
XX
CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides (AAY40601-
CC Y40609) together form a single-chain scaffold protein which contains at
CC least 1 disulfide bond, contains less than 10% alpha helix and contains
CC at least 6 beta-strands. The scaffold protein is constructed of beta
CC strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens

```
GN   LE-ACS1B.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97351561; PubMed=9207843;
RA   Oetiker J.H., Olson D.C., Shiu O.Y., Yang S.F.;
RT   "Differential induction of seven 1-aminocyclopropane-1-carboxylate
RT   synthase genes by elicitor in suspension cultures of tomato
RT   (Lycopersicon esculentum).";
RL   Plant Mol. Biol. 34:275-286(1997).
DR   EMBL; U75692; AAC49682.1; -.
DR   GO; GO:0016847; F:1-aminocyclopropane-1-carboxylate synthase . . .; IEA.
DR   GO; GO:0016829; F:lyase activity; IEA.
KW   Lyase.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  828 MW;  71B412C7377415D0 CRC64;

  Query Match             27.3%;  Score 9;  DB 10;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 SR 7
              ||
Db          1 SR 2


RESULT 14
Q8K3H6
ID   Q8K3H6      PRELIMINARY;      PRT;     7 AA.
AC   Q8K3H6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Collagenase-3 (Fragment).
GN   MMP13.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   Prietto S.M., Lyons J.G.;
RT   "Intron 1 of Rattus norvegicus MMP13.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY101357; AAM51172.1; -.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  907 MW;  63373B51EB1DD9A0 CRC64;

  Query Match             27.3%;  Score 9;  DB 11;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
              ||
Db          6 LK 7


RESULT 15
Q63480
ID   Q63480      PRELIMINARY;      PRT;     7 AA.
AC   Q63480;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   TR4-NS orphan receptor (Fragment).
GN   TR4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96198747; PubMed=8612486;
RA   Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,
RA   Detera-Wadleigh S.D.;
RT   "Splice variants of rat TR4 orphan receptor: differential expression
RT   of novel sequences in the 5'-untranslated region and C-terminal
RT   domain.";
RL   Endocrinology 137:1562-1571(1996).
DR   EMBL; U59125; AAB02827.1; -.
DR   GO; GO:0004872; F:receptor activity; IEA.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  758 MW;  672AA87864005350 CRC64;

  Query Match             27.3%;  Score 9;  DB 11;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 KI 5
              ||
Db          1 KI 2


Search completed: June  1, 2004, 11:18:37
Job time : 32.3333 secs
```

```
AC   Q8GL12;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   PF-50 protein (Fragment).
GN   PF-50.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid group cp32-9.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N40;
RA   Stevenson B., Miller J.C.;
RT   "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
RT   prophages: conservation amidst diversity.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY142100; AAN17911.1; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW   Plasmid.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  849 MW;  6337244330569ED0 CRC64;

  Query Match             27.3%;  Score 9;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
               ||
Db          6 LK 7


RESULT 10
Q8GL04
ID   Q8GL04      PRELIMINARY;      PRT;     7 AA.
AC   Q8GL04;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   PF-50 protein (Fragment).
GN   PF-50.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid group cp32-5.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=93-0107;
RA   Stevenson B., Miller J.C.;
RT   "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
RT   prophages: conservation amidst diversity.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY142103; AAN17848.1; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW   Plasmid.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  914 MW;  6337244330504310 CRC64;

  Query Match             27.3%;  Score 9;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
               ||
Db          6 LK 7


RESULT 11
Q8GL00
ID   Q8GL00      PRELIMINARY;      PRT;     7 AA.
AC   Q8GL00;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   PF-50 protein (Fragment).
GN   PF-50.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid group cp32-13.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CA15;
RA   Stevenson B., Miller J.C.;
RT   "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
RT   prophages: conservation amidst diversity.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY142106; AAN17857.1; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW   Plasmid.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  928 MW;  6337233050437350 CRC64;

  Query Match             27.3%;  Score 9;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
               ||
Db          6 LK 7


RESULT 12
Q15897
ID   Q15897      PRELIMINARY;      PRT;     7 AA.
AC   Q15897;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   (Clone XP6A11A) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA   Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA   Caskey C.T.H.;
RT   "Isolation of chromosome-specific genes by reciprocal probing of
RT   arrayed cDNAs and cosmid libraries.";
RL   Hum. Mol. Genet. 0:0-0(1995).
DR   EMBL; L32077; AAA73887.1; -.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  814 MW;  672B1DD3372046B0 CRC64;

  Query Match             27.3%;  Score 9;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
               ||
Db          3 LK 4


RESULT 13
P93233
ID   P93233      PRELIMINARY;      PRT;     7 AA.
AC   P93233;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14)
DE   (Fragment).
```

```
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D273-10B;
RX   MEDLINE=81069885; PubMed=6254986;
RA   Bonitz S.G., Coruzzi G., Thalenfeld B., Tzagoloff A., Macino G.;
RT   "Assembly of the mitochondrial membrane system: Structure and
RT   nucleotide sequence of the gene coding for subunit 1 of yeast
RT   cytochrome oxidase.";
RL   J. Biol. Chem. 255:11927-11941(1980).
DR   EMBL; V00694; CAA24066.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  859 MW;  75B7232362CDC460 CRC64;

  Query Match             33.3%;  Score 11;  DB 8;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 KIS 6
              |:|
Db          5 KLS 7


RESULT 6
P82541
ID   P82541      PRELIMINARY;      PRT;     6 AA.
AC   P82541;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 30S ribosomal protein S19 beta (Fragment).
OS   Spinacia oleracea (Spinach).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435797; PubMed=10874039;
RA   Yamaguchi K., von Knoblauch K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the small subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 37:28455-28465(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.
CC   -!- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.
CC   -!- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA
CC       FORM IS THE MINOR BASIC FORM.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.
CC   -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002222; Ribosomal_S19.
DR   PROSITE; PS00323; RIBOSOMAL_S19; PARTIAL.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  732 MW;  63333735A411C000 CRC64;

  Query Match             30.3%;  Score 10;  DB 10;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLK 4
              :||
Db          3 SLK 5


RESULT 7
Q08433
ID   Q08433      PRELIMINARY;      PRT;     4 AA.
AC   Q08433;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Bilirubin UDP-glucuronosyltransferase (Fragment).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Gunn;
RX   MEDLINE=91282758; PubMed=1840486;
RA   Sato H., Aono S., Kashiwamata S., Koiwai O.;
RT   "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
RT   hyperbilirubinemic Gunn rat.";
RL   Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
DR   EMBL; S38636; AAB19259.1; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
KW   Transferase.
FT   NON_TER       1      1
SQ   SEQUENCE   4 AA;  473 MW;  633732C420000000 CRC64;

  Query Match             27.3%;  Score 9;  DB 11;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
              ||
Db          3 LK 4


RESULT 8
O07354
ID   O07354      PRELIMINARY;      PRT;     7 AA.
AC   O07354;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NifK (Fragment).
GN   NIFK.
OS   Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothece PCC 8801).
OC   Bacteria; Cyanobacteria; Chroococcales; Cyanothece.
OX   NCBI_TaxID=41431;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RF-1;
RX   MEDLINE=99231861; PubMed=10217509;
RA   Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RT   "Organization and expression of nitrogen-fixation genes in the aerobic
RT   nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
RT   RF-1.";
RL   Microbiology 145:743-753(1999).
DR   EMBL; AF003700; AAC35193.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  849 MW;  7412C72AA9D5B030 CRC64;

  Query Match             27.3%;  Score 9;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          3 FDL 5


RESULT 9
Q8GL12
ID   Q8GL12      PRELIMINARY;      PRT;     7 AA.
```

```
Qy          5 ISR 7
              |||
Db          2 ISR 4


RESULT 2
P82182
ID   P82182      PRELIMINARY;      PRT;     6 AA.
AC   P82182;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 50S ribosomal protein L10 gamma (Fragment).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435798; PubMed=10874046;
RA   Yamaguchi K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the 50 S subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 275:28466-28482(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC   -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002363; Ribosomal_L10eub.
DR   PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  675 MW;  6321B415B05DB000 CRC64;

  Query Match             39.4%;  Score 13;  DB 10;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ISR 7
              |||
Db          2 ISR 4


RESULT 3
Q7Z1C0
ID   Q7Z1C0      PRELIMINARY;      PRT;     7 AA.
AC   Q7Z1C0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein W01B11.6.
GN   W01B11.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Wilson R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Bradshaw H., Graves T., Blair T.;
RT   "The sequence of C. elegans cosmid W01B11.";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Wilson R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF043704; AAO38592.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   7 AA;  874 MW;  72D1A9DB5041A6F0 CRC64;

  Query Match             36.4%;  Score 12;  DB 5;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              ||:
Db          5 FTV 7


RESULT 4
P83530
ID   P83530      PRELIMINARY;      PRT;     7 AA.
AC   P83530;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Unknown protein from 2D-page (Fragment).
OS   Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1625;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=DSM 20451;
RX   PubMed=12112860;
RA   Drews O., Weiss W., Reil G., Parlar H., Wait R., Goerg A.;
RT   "High pressure effects step-wise altered protein expression in
RT   Lactobacillus sanfranciscensis.";
RL   Proteomics 2:765-774(2002).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS UNKNOWN
CC       PROTEIN IS: 15 KDA.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  790 MW;  6AAB02CAA731B2A0 CRC64;

  Query Match             33.3%;  Score 11;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKI 5
              || :
Db          2 TLDV 5


RESULT 5
Q95945
ID   Q95945      PRELIMINARY;      PRT;     7 AA.
AC   Q95945;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Inside intron 5 (Fragment).
OS   Saccharomyces cerevisiae (Baker's yeast).
OG   Mitochondrion.
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:53:46 ; Search time 31.3333 Seconds
(without alignments)
70.488 Million cell updates/sec

Title: US-09-712-819C-6
Perfect score: 33
Sequence: 1 FTLKISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 74

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	13	39.4	6	10	P82181	P82181 spinacia ol
2	13	39.4	6	10	P82182	P82182 spinacia ol
3	12	36.4	7	5	Q7Z1C0	Q7z1c0 caenorhabdi
4	11	33.3	7	2	P83530	P83530 lactobacill
5	11	33.3	7	8	Q95945	Q95945 saccharomyc
6	10	30.3	6	10	P82541	P82541 spinacia ol
7	9	27.3	4	11	Q08433	Q08433 rattus sp.
8	9	27.3	7	2	O07354	O07354 synechococc
9	9	27.3	7	2	Q8GL12	Q8gl12 borrelia bu
10	9	27.3	7	2	Q8GL04	Q8gl04 borrelia bu
11	9	27.3	7	2	Q8GL00	Q8gl00 borrelia bu
12	9	27.3	7	4	Q15897	Q15897 homo sapien
13	9	27.3	7	10	P93233	P93233 lycopersico
14	9	27.3	7	11	Q8K3H6	Q8k3h6 rattus norv
15	9	27.3	7	11	Q63480	Q63480 rattus norv
16	9	27.3	7	11	O55184	O55184 rattus norv
17	9	27.3	7	15	Q8JE81	Q8je81 human immun
18	8	24.2	7	2	P70804	P70804 azotobacter
19	8	24.2	7	2	Q47029	Q47029 enterobacte
20	8	24.2	7	12	Q9YVE3	Q9yve3 human adeno
21	8	24.2	7	12	Q9YIQ9	Q9yiq9 human adeno
22	8	24.2	7	12	Q9YIR0	Q9yir0 human adeno
23	7	21.2	5	2	P83073	P83073 bacillus ce
24	7	21.2	7	2	P72081	P72081 nocardia la
25	7	21.2	7	4	Q8NHH7	Q8nhh7 homo sapien
26	7	21.2	7	12	Q66205	Q66205 transmissib
27	7	21.2	7	12	Q9YQ10	Q9yq10 transmissib
28	6	18.2	5	13	P83308	P83308 gallus gall
29	6	18.2	6	2	P83533	P83533 lactobacill
30	6	18.2	7	2	O50556	O50556 actinobacil
31	6	18.2	7	2	Q54248	Q54248 streptomyce
32	6	18.2	7	2	Q8KMS9	Q8kms9 enterobacte
33	6	18.2	7	2	O34028	O34028 sphingomona
34	6	18.2	7	4	Q15903	Q15903 homo sapien
35	6	18.2	7	6	Q28742	Q28742 oryctolagus
36	6	18.2	7	8	P92214	P92214 amblyopyrum
37	6	18.2	7	8	P92393	P92393 hordeum vul
38	6	18.2	7	8	P92403	P92403 lophopyrum
39	6	18.2	7	8	P92427	P92427 peridictyon
40	6	18.2	7	8	P92430	P92430 aegilops ta
41	6	18.2	7	8	P92221	P92221 bromus iner
42	6	18.2	7	8	O98866	O98866 spinacia ol
43	6	18.2	7	8	P92425	P92425 pseudoroegn
44	6	18.2	7	8	P92381	P92381 hordeum bra
45	6	18.2	7	8	P92387	P92387 henrardia p

ALIGNMENTS

RESULT 1
P82181
ID P82181 PRELIMINARY; PRT; 6 AA.
AC P82181;
DT 01-JUN-2000 (TrEMBLrel. 14, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chloroplast 50S ribosomal protein L10 beta (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. ALWARO; TISSUE=Leaf;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0019843; F:rRNA binding; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR InterPro; IPR002363; Ribosomal_L10eub.
DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 39.4%; Score 13; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Multigene family.
SQ   SEQUENCE   7 AA;  782 MW;  672879CDCB476AC0 CRC64;

  Query Match             24.2%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          5 FGL 7


RESULT 15
ALL5_CARMA
ID   ALL5_CARMA     STANDARD;      PRT;     7 AA.
AC   P81808;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 5.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       7       7       AMIDATION.
SQ   SEQUENCE   7 AA;  781 MW;  672879CDCB476420 CRC64;

  Query Match             24.2%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          5 FGL 7


Search completed: June  1, 2004, 11:16:49
Job time : 7.66667 secs
```

```
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       5       5       AMIDATION (POTENTIAL).
SQ   SEQUENCE   5 AA;  586 MW;  672879D5AB300000 CRC64;

  Query Match             24.2%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          3 FGL 5


RESULT 11
PSK_DAUCA
ID   PSK_DAUCA      STANDARD;      PRT;     5 AA.
AC   P58261;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Phytosulfokine-alpha (PSK-alpha) [Contains: Phytosulfokine-beta (PSK-
DE   beta)].
OS   Daucus carota (Carrot).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   campanulids; Apiales; Apiaceae; Apioideae; Scandiceae; Daucinae;
OC   Daucus.
OX   NCBI_TaxID=4039;
RN   [1]
RP   SEQUENCE, AND IDENTIFICATION BY MASS SPECTROMETRY.
RC   STRAIN=cv. US-Harumakigosun;
RX   MEDLINE=20212743; PubMed=10750705;
RA   Hanai H., Matsuno T., Yamamoto M., Matsubayashi Y., Kobayashi T.,
RA   Kamada H., Sakagami Y.;
RT   "A secreted peptide growth factor, phytosulfokine, acting as a
RT   stimulatory factor of carrot somatic embryo formation.";
RL   Plant Cell Physiol. 41:27-32(2000).
CC   -!- FUNCTION: In presence of 2,4-D, stimulates proliferation of the
CC       cells, but does not stimulate differentiation into the somatic
CC       embryos.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: Sulfation is important for activity and for the binding to a
CC       putative membrane receptor (By similarity).
CC   -!- SIMILARITY: Belongs to the phytosulfokine family.
KW   Growth factor; Sulfation.
FT   PEPTIDE       1       4       PHYTOSULFOKINE-BETA.
FT   MOD_RES       1       1       SULFATION.
FT   MOD_RES       3       3       SULFATION.
SQ   SEQUENCE   5 AA;  687 MW;  76C1BB504B300000 CRC64;

  Query Match             24.2%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FT 2
              :|
Db          3 YT 4


RESULT 12
ALL2_CARMA
ID   ALL2_CARMA     STANDARD;      PRT;     7 AA.
AC   P81805;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 2.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       7       7       AMIDATION (POTENTIAL).
SQ   SEQUENCE   7 AA;  770 MW;  672879CDCB5DDB70 CRC64;

  Query Match             24.2%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          5 FGL 7


RESULT 13
ALL3_CARMA
ID   ALL3_CARMA     STANDARD;      PRT;     7 AA.
AC   P81806;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 3.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Multigene family.
SQ   SEQUENCE   7 AA;  796 MW;  672879CDCB476B70 CRC64;

  Query Match             24.2%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          5 FGL 7


RESULT 14
ALL4_CARMA
ID   ALL4_CARMA     STANDARD;      PRT;     7 AA.
AC   P81807;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 4.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
```

```
RL   J. Biol. Chem. 263:14574-14578(1988).
CC   -!- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       HEMOLYSIN PLASMID PCF10.
DR   PIR; A30812; A30812.
KW   Pheromone.
SQ   SEQUENCE   7 AA;  790 MW;  72C9D2C731B2C740 CRC64;

  Query Match             27.3%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TL 3
              ||
Db          3 TL 4


RESULT 7
CHOX_ALCSP
ID   CHOX_ALCSP     STANDARD;      PRT;     7 AA.
AC   P16101;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-APR-1990 (Rel. 14, Last annotation update)
DE   Choline oxidase (EC 1.1.3.17) (Fragment).
OS   Alcaligenes sp.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Alcaligenes.
OX   NCBI_TaxID=512;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=81006769; PubMed=6997283;
RA   Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;
RT   "Identification and properties of the prosthetic group of choline
RT   oxidase from Alcaligenes sp.";
RL   J. Biochem. 88:197-203(1980).
CC   -!- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).
DR   PIR; A15398; A15398.
KW   Oxidoreductase.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  839 MW;  7415B1E457644AC0 CRC64;

  Query Match             27.3%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 SR 7
              ||
Db          6 SR 7


RESULT 8
CIA_ENTFA
ID   CIA_ENTFA      STANDARD;      PRT;     7 AA.
AC   P11932;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cAM373 (Clumping-inducing agent) (CIA).
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=87005252; PubMed=3093276;
RA   Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
RA   White B.A., An F.Y., Clewell D.B., Suzuki A.;
RT   "Isolation and structure of the Streptococcus faecalis sex pheromone,
RT   cAM373.";
RL   FEBS Lett. 206:69-72(1986).
CC   -!- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
CC       HARBORING PAM373.
CC   -!- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
CC       SPECIFICITY OF PHEROMONES TO PLASMIDS.
CC   -!- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
DR   PIR; A25269; A25269.
KW   Pheromone.
SQ   SEQUENCE   7 AA;  734 MW;  75BDD72059C05DB0 CRC64;

  Query Match             27.3%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          3 FIL 5


RESULT 9
UC24_MAIZE
ID   UC24_MAIZE     STANDARD;      PRT;     7 AA.
AC   P80630;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Unknown protein from 2D-page of etiolated coleoptile (Spot 447)
DE   (Fragment).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Coleoptile;
RA   Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA   Pernollet J.-C., Zivy M., de Vienne D.;
RT   "The maize two dimensional gel protein database: towards an integrated
RT   genome analysis program.";
RL   Theor. Appl. Genet. 93:997-1005(1996).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 6.0, its MW is: 30.0 kDa.
DR   Maize-2DPAGE; P80630; COLEOPTILE.
DR   MaizeDB; 123956; -.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  665 MW;  6DC1B5B33DC1B5D0 CRC64;

  Query Match             27.3%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLK 4
              | |
Db          2 TAK 4


RESULT 10
AL14_CARMA
ID   AL14_CARMA     STANDARD;      PRT;     5 AA.
AC   P81817;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 14.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
```

```
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litori electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:639-645(1999).
CC   -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
KW   Amphibian defense peptide.
SQ   SEQUENCE   5 AA;  570 MW;  71A9C9C862A00000 CRC64;

  Query Match             33.3%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FT 2
              ||
Db          4 FT 5


RESULT 3
GFRP_MOUSE
ID   GFRP_MOUSE     STANDARD;      PRT;     7 AA.
AC   P99025;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
GN   GCHFR OR GFRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Liver;
RA   Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
RA   Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA   Cowthorne M.;
RL   Submitted (AUG-1998) to Swiss-Prot.
CC   -!- FUNCTION: Mediates tetrahydrobiopterin inhibition of GTP
CC       cyclohydrolase I. This inhibition is reversed by L-phenylalanine
CC       (By similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
DR   SWISS-2DPAGE; P99025; MOUSE.
FT   INIT_MET      0      0
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  806 MW;  71B5B057273B4700 CRC64;

  Query Match             30.3%;  Score 10;  DB 1;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 1.4e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LKIS 6
              | ||
Db          3 LLIS 6


RESULT 4
UF01_MOUSE
ID   UF01_MOUSE     STANDARD;      PRT;     5 AA.
AC   P38639;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Unknown protein from 2D-page of fibroblasts (P19) (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Fibroblast;
RX   MEDLINE=95009907; PubMed=7523108;
RA   Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT   "Separation and sequencing of familiar and novel murine proteins
RT   using preparative two-dimensional gel electrophoresis.";
RL   Electrophoresis 15:735-745(1994).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 6.6, its MW is: 19 kDa.
FT   NON_TER       5      5
SQ   SEQUENCE   5 AA;  717 MW;  7364087043100000 CRC64;

  Query Match             27.3%;  Score 9;  DB 1;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ISR 7
              | |
Db          2 IGR 4


RESULT 5
UN06_CLOPA
ID   UN06_CLOPA     STANDARD;      PRT;     6 AA.
AC   P81351;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Unknown protein CP 6 from 2D-page (Fragment).
OS   Clostridium pasteurianum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1501;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=W5;
RX   MEDLINE=98291870; PubMed=9629918;
RA   Flengsrud R., Skjeldal L.;
RT   "Two-dimensional gel electrophoresis separation and N-terminal
RT   sequence analysis of proteins from Clostridium pasteurianum W5.";
RL   Electrophoresis 19:802-806(1998).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 5.0, ITS MW IS: 75.9 kDa.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  657 MW;  605B1DC1A45A8000 CRC64;

  Query Match             27.3%;  Score 9;  DB 1;  Length 6;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKI 5
              | :|
Db          3 TAEI 6


RESULT 6
CCF1_ENTFA
ID   CCF1_ENTFA     STANDARD;      PRT;     7 AA.
AC   P20104;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cCF10.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=89008313; PubMed=3139658;
RA   Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,
RA   Adsit J.C., Dunny G.M., Suzuki A.;
RT   "Structure of cCF10, a peptide sex pheromone which induces
RT   conjugative transfer of the Streptococcus faecalis tetracycline
RT   resistance plasmid, pCF10.";
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:41:50 ; Search time 7.66667 Seconds
(without alignments)
47.542 Million cell updates/sec

Title: US-09-712-819C-6
Perfect score: 33
Sequence: 1 FTLKISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 88

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	11	33.3	5	1	RE31_LITRU	P82072 litoria rub
2	11	33.3	5	1	RE32_LITRU	P82073 litoria rub
3	10	30.3	7	1	GFRP_MOUSE	P99025 mus musculu
4	9	27.3	5	1	UF01_MOUSE	P38639 mus musculu
5	9	27.3	6	1	UN06_CLOPA	P81351 clostridium
6	9	27.3	7	1	CCF1_ENTFA	P20104 enterococcu
7	9	27.3	7	1	CHOX_ALCSP	P16101 alcaligenes
8	9	27.3	7	1	CIA_ENTFA	P11932 enterococcu
9	9	27.3	7	1	UC24_MAIZE	P80630 zea mays (m
10	8	24.2	5	1	AL14_CARMA	P81817 carcinus ma
11	8	24.2	5	1	PSK_DAUCA	P58261 daucus caro
12	8	24.2	7	1	ALL2_CARMA	P81805 carcinus ma
13	8	24.2	7	1	ALL3_CARMA	P81806 carcinus ma
14	8	24.2	7	1	ALL4_CARMA	P81807 carcinus ma
15	8	24.2	7	1	ALL5_CARMA	P81808 carcinus ma
16	8	24.2	7	1	ALL7_CYDPO	P82158 cydia pomon
17	8	24.2	7	1	CARP_MYTED	P10420 mytilus edu
18	8	24.2	7	1	FAR5_HIRME	P42564 hirudo medi
19	7	21.2	3	1	LUXE_VIBFI	P24272 vibrio fisc
20	7	21.2	6	1	LOK1_LOCMI	P41491 locusta mig
21	7	21.2	6	1	VP19_HSV1K	P23210 herpes simp
22	6	18.2	4	1	ACH1_ACHFU	P35904 achatina fu
23	6	18.2	4	1	FAR3_HIRME	P42562 hirudo medi
24	6	18.2	4	1	FAR4_HIRME	P42563 hirudo medi
25	6	18.2	4	1	FFKA_ANTEL	P58705 anthopleura
26	6	18.2	4	1	FLRF_HIRME	P42561 hirudo medi
27	6	18.2	4	1	FLRN_ANTEL	P58707 anthopleura
28	6	18.2	4	1	FMRF_MACNI	P01162 macrocallis
29	6	18.2	4	1	FYRI_ANTEL	P58706 anthopleura
30	6	18.2	4	1	OCP1_OCTMI	P58648 octopus min
31	6	18.2	5	1	EI03_LITRU	P82099 litoria rub
32	6	18.2	5	1	EI04_LITRU	P82100 litoria rub
33	6	18.2	5	1	FARP_ARTTR	P41853 artioposthi
34	6	18.2	5	1	PAP2_PARMA	P81864 pardachirus
35	6	18.2	5	1	RE11_LITRU	P82070 litoria rub
36	6	18.2	5	1	RE21_LITRU	P82071 litoria rub
37	6	18.2	5	1	SUGA_ACHDO	P19991 acheta dome
38	6	18.2	5	1	TPIS_CANFA	P54714 canis famil
39	6	18.2	5	1	UC22_MAIZE	P80628 zea mays (m
40	6	18.2	6	1	ACPH_RABIT	P25154 oryctolagus
41	6	18.2	6	1	CIP1_MYTED	P13736 mytilus edu
42	6	18.2	6	1	CIP2_MYTED	P13737 mytilus edu
43	6	18.2	6	1	EI01_LITRU	P82096 litoria rub
44	6	18.2	6	1	FARP_MONEX	P41966 moniezia ex
45	6	18.2	7	1	EI05_LITRU	P82101 litoria rub

ALIGNMENTS

RESULT 1
RE31_LITRU
ID RE31_LITRU STANDARD; PRT; 5 AA.
AC P82072;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rubellidin 3.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC Pelodryadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Skin secretion;
RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australin red tree frog
RT 'Litoria rubella'. The skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
CC activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC -!- MASS SPECTROMETRY: MW=655; METHOD=FAB.
KW Amphibian defense peptide; Amidation.
FT MOD_RES 5 5 AMIDATION.
SQ SEQUENCE 5 AA; 656 MW; 71A9C9CB10300000 CRC64;

Query Match 33.3%; Score 11; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FT 2
||
Db 4 FT 5

RESULT 2
RE32_LITRU
ID RE32_LITRU STANDARD; PRT; 5 AA.
AC P82073;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rubellidin 3.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC Pelodryadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;

T-cell receptor beta chain V-D-J region (141-1BC) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0577; PT0574
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0577
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1BC
A;Accession: PT0574
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FE2>
A;Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1Q
C;Keywords: T-cell receptor

  Query Match 27.3%; Score 9; DB 2; Length 5;
  Best Local Similarity 100.0%; Pred. No. 2.8e+05;
  Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 SR 7
   ||
Db 3 SR 4

RESULT 15
PT0565
T-cell receptor beta chain V-D-J region (141-1CF) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0565
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0565
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match 27.3%; Score 9; DB 2; Length 5;
  Best Local Similarity 100.0%; Pred. No. 2.8e+05;
  Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 SR 7
   ||
Db 3 SR 4

Search completed: June 1, 2004, 11:19:23
Job time : 11.6667 secs

```
A;Cross-references: EMBL:U31309; NID:g974285; PID:g974292
A;Experimental source: strain s6PT2xs6PT3; 8 month seedlings

  Query Match             27.3%;  Score 9;  DB 2;  Length 4;
  Best Local Similarity   33.3%;  Pred. No. 2.8e+05;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LKI 5
              :|:
Db          1 MKL 3


RESULT 9
A60521
glycogen phosphorylase (EC 2.4.1.1), muscle - mullet (Liza ramada) (fragment)
N;Alternate names: glycogen phosphorylase b
C;Species: Liza ramada
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2003
C;Accession: A60521
R;Bonamusa, L.; Baanante, I.V.
Comp. Biochem. Physiol. B 95, 295-301, 1990
A;Title: Purification and characterization of glycogen phosphorylase B from skeletal mus
A;Reference number: A60521; MUID:90227907; PMID:2109669
A;Accession: A60521
A;Molecule type: protein
A;Residues: 1-5 <BON>
C;Superfamily: glucan phosphorylase
C;Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein
F;3/Binding site: phosphate (Ser) (covalent) (by phosphorylase b kinase) #status experim

  Query Match             27.3%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 KIS 6
              :||
Db          1 QIS 3


RESULT 10
E42364
flagellar protein fliR - Salmonella typhimurium (fragment)
C;Species: Salmonella typhimurium
C;Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 30-Sep-1993
C;Accession: E42364
R;Vogler, A.P.; Homma, M.; Irikura, V.M.; Macnab, R.M.
J. Bacteriol. 173, 3564-3572, 1991
A;Title: Salmonella typhimurium mutants defective in flagellar filament regrowth and seq
A;Reference number: A42364; MUID:91258342; PMID:1646201
A;Accession: E42364
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-5 <VOG>
A;Cross-references: GB:M62408

  Query Match             27.3%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TL 3
              ||
Db          3 TL 4


RESULT 11
A44955
alkanal monooxygenase (FMN-linked) (EC 1.14.14.3) alpha chain - Vibrio harveyi (fragment
C;Species: Vibrio harveyi
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-May-2000
C;Accession: A44955
R;Paquatte, O.; Tu, S.C.
Photochem. Photobiol. 50, 817-825, 1989
A;Title: Chemical modification and characterization of the alpha cysteine 106 at the Vib
A;Reference number: A44955; MUID:90175700; PMID:2626493
A;Accession: A44955
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-5 <PAQ>
C;Keywords: FMN; luminescence; monooxygenase; oxidoreductase

  Query Match             27.3%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ISR 7
              | |
Db          3 IXR 5


RESULT 12
S11127
phosphoprotein, bone - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 04-Mar-2000
C;Accession: S11127; S11128
R;Mikuni-Takagaki, Y.; Glimcher, M.J.
Biochem. J. 268, 585-591, 1990
A;Title: Post-translational processing of chicken bone phosphoproteins. Identifica
A;Reference number: S11127; MUID:90303246; PMID:2363696
A;Accession: S11127
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-5 <MIK1>
A;Accession: S11128
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'X',2-5 <MIK2>
C;Keywords: phosphoprotein

  Query Match             27.3%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity   33.3%;  Pred. No. 2.8e+05;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ISR 7
              :|:
Db          3 VSK 5


RESULT 13
PT0525
T-cell receptor beta chain V-D-J region (100-4J) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0525
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regio
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0525
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             27.3%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 SR 7
              ||
Db          2 SR 3


RESULT 14
PT0577
```

T14910
hypothetical protein - parsley
C;Species: Petroselinum crispum (parsley)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14910
R;Kircher, S.; Ledger, S.; Hayashi, H.; Weisshaar, B.; Schafer, E.; Frohnmeyer, H.
Mol. Gen. Genet. 257, 595-605, 1998
A;Title: CPRF4a, a novel plant bZIP protein of the CPRF family: comparative analysis of
A;Reference number: Z18261; MUID:98265918; PMID:9604882
A;Accession: T14910
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-5 <KIR>
A;Cross-references: EMBL:Y10810; NID:g3336904; PIDN:CAA71769.1; PID:g3336905
A;Experimental source: ssp. Hamburger Schnitt

Query Match 36.4%; Score 12; DB 2; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ISR 7
:||
Db 2 VSR 4

RESULT 4
PT0644
T-cell receptor beta chain V-D-J region (111-1G) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0644
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0644
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 33.3%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FT 2
||
Db 4 FT 5

RESULT 5
PS0254
18K protein 5507 - rice (strain Nihonbare) (fragment)
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Mar-1995
C;Accession: PS0254
R;Tsugita, A.
submitted to JIPID, April 1993
A;Reference number: PS0206
A;Accession: PS0254
A;Molecule type: protein
A;Residues: 1-7 <TSU>
A;Experimental source: leaf, chloroplast, strain Nihonbare
A;Note: molecular weight 18K, pI 4.4

Query Match 30.3%; Score 10; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LKISR 7
| |::
Db 1 LAIAK 5

RESULT 6
T13892
cytochrome-c oxidase (EC 1.9.3.1) chain I [imported] - river lamprey mitochondrio
C;Species: mitochondrion Lampetra fluviatilis (river lamprey)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: T13892
R;Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A;Title: The main features of the craniate mitochondrial DNA between the ND1 and
A;Reference number: Z17775; MUID:97398704; PMID:9254918
A;Accession: T13892
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3 <DEL>
A;Cross-references: EMBL:Y09528; NID:g2340016; PIDN:CAA70721.1; PID:g4379123
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

Query Match 27.3%; Score 9; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TL 3
||
Db 2 TL 3

RESULT 7
I40804
endoglucanase F - Clostridium thermocellum (fragment)
C;Species: Clostridium thermocellum
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Aug-1996
C;Accession: I40804
R;Mishra, S.; Beguin, P.; Aubert, J.
J. Bacteriol. 173, 80-85, 1991
A;Title: Transcription of clostridium thermoncellum endoglucanase genes celF and
A;Reference number: I40804; MUID:91100322; PMID:1987137
A;Accession: I40804
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-4 <RES>
A;Cross-references: GB:M64363; NID:g144771
C;Genetics:
A;Gene: celF
A;Start codon: TTG

Query Match 27.3%; Score 9; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KI 5
||
Db 3 KI 4

RESULT 8
T46627
hypothetical protein c4 - loblolly pine
C;Species: Pinus taeda (loblolly pine)
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
C;Accession: T46627
R;Chang, S.; Puryea, J.; Funkhouser, E.A.; Newton, R.J.; Cairney, J.
submitted to the EMBL Data Library, July 1995
A;Description: Cloning of a chitinase homolog which lacks chitin binding sites and
A;Reference number: Z23105
A;Accession: T46627
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-4 <CHA>

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:56:46 ; Search time 11.6667 Seconds
(without alignments)
57.715 Million cell updates/sec

Title: US-09-712-819C-6
Perfect score: 33
Sequence: 1 FTLKISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 457

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	15	45.5	7	2	S19630	ribosomal protein
2	13	39.4	7	2	A28709	phosphonoacetaldeh
3	12	36.4	5	2	T14910	hypothetical prote
4	11	33.3	5	2	PT0644	T-cell receptor be
5	10	30.3	7	2	PS0254	18K protein 5507 -
6	9	27.3	3	3	T13892	cytochrome-c oxida
7	9	27.3	4	2	I40804	endoglucanase F -
8	9	27.3	4	2	T46627	hypothetical prote
9	9	27.3	5	2	A60521	glycogen phosphory
10	9	27.3	5	2	E42364	flagellar protein
11	9	27.3	5	2	A44955	alkanal monooxygen
12	9	27.3	5	2	S11127	phosphoprotein, bo
13	9	27.3	5	2	PT0525	T-cell receptor be
14	9	27.3	5	2	PT0577	T-cell receptor be
15	9	27.3	5	2	PT0565	T-cell receptor be
16	9	27.3	5	2	PT0700	T-cell receptor be
17	9	27.3	5	2	S69237	surface protein te
18	9	27.3	6	2	A60986	N-formyl oligopept
19	9	27.3	6	2	A43766	28K ubiquitin-immu
20	9	27.3	6	2	I37263	Y protein - human
21	9	27.3	6	2	B26206	alpha-1,4-glucan-p
22	9	27.3	6	2	I65546	MHC H2-L antigen -
23	9	27.3	6	2	PT0518	T-cell receptor be
24	9	27.3	6	2	PT0662	T-cell receptor be
25	9	27.3	6	2	I49424	cytotoxic T-lympho
26	9	27.3	7	2	JN0859	peptidyl-dipeptida
27	9	27.3	7	2	A15398	choline oxidase (E
28	9	27.3	7	2	B39127	phosphotransferase
29	9	27.3	7	2	S25266	pilE protein - Esc
30	9	27.3	7	2	A25269	sex pheromone cAM3
31	9	27.3	7	2	A30812	sex pheromone cCF1
32	9	27.3	7	2	PN0150	omega-gliadine 1'
33	9	27.3	7	2	S78024	ribosomal protein
34	9	27.3	7	2	E48394	glycoprotein compo
35	9	27.3	7	2	I48086	DNA topoisomerase
36	9	27.3	7	2	PT0671	T-cell receptor be
37	9	27.3	7	2	S66442	glutathione S-tran
38	9	27.3	7	2	B48394	major fat-globule
39	9	27.3	7	2	PN0649	pullulanase (EC 3.
40	9	27.3	7	2	S09066	globulin IV alpha
41	8	24.2	4	2	I40505	hypothetical prote
42	8	24.2	5	2	I39964	ribosomal protein
43	8	24.2	5	2	I39966	ribosomal protein
44	8	24.2	5	2	I39965	ribosomal protein
45	8	24.2	5	2	B22565	R-phycoerythrin al

ALIGNMENTS

RESULT 1
S19630
ribosomal protein L30 - Streptomyces griseus (fragment)
C;Species: Streptomyces griseus
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Jun-1997
C;Accession: S19630
R;Ochi, K.
Int. J. Syst. Bacteriol. 42, 144-150, 1992
A;Title: Electrophoretic heterogeneity of ribosomal protein AT-L30 among actinomyce
A;Reference number: S19630; MUID:92144363; PMID:1736962
A;Accession: S19630
A;Molecule type: protein
A;Residues: 1-7 <OCH>
A;Experimental source: strain IFO 13189
C;Superfamily: Escherichia coli ribosomal protein L30
C;Keywords: protein biosynthesis; ribosome

Query Match 45.5%; Score 15; DB 2; Length 7;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 LKISR 7
              |||::
Db          3 LKITQ 7
```

RESULT 2
A28709
phosphonoacetaldehyde hydrolase - Bacillus cereus (fragment)
C;Species: Bacillus cereus
C;Date: 22-Aug-1988 #sequence_revision 22-Aug-1988 #text_change 30-Sep-1993
C;Accession: A28709
R;Olsen, D.B.; Hepburn, T.W.; Moos, M.; Mariano, P.S.; Dunaway-Mariano, D.
Biochemistry 27, 2229-2234, 1988
A;Title: Investigation of the Bacillus cereus phosphonoacetaldehyde hydrolase. Evic
idue.
A;Reference number: A28709; MUID:88241058; PMID:3132206
A;Accession: A28709
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <OLS>

Query Match 39.4%; Score 13; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 LKI 5
              |||
Db          1 LKI 3
```

RESULT 3

; FILING DATE: 10/14/93
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5459063e
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-136-743B-30

Query Match 48.6%; Score 17; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TDF 3
              |||
Db          5 TDF 7
```

RESULT 15
US-08-136-743B-31
; Sequence 31, Application US/08136743B
; Patent No. 5459063
; GENERAL INFORMATION:
; APPLICANT: Barry S. Cooperman, Harvey Rubin,
; APPLICANT: Jerome Salem, and Alison L. Fisher
; TITLE OF INVENTION: "Plasmodium falciparum Ribonu-
; TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide inhibit
; TITLE OF INVENTION: Thereof"
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The University of Pennsylvania
; STREET: Suite 330
; STREET: 3700 Market Street
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19104-3246
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136,743B
; FILING DATE: 10/14/93
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5459063e
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-136-743B-31

Query Match 48.6%; Score 17; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TDF 3
              |||
Db          5 TDF 7
```

Search completed: June 1, 2004, 11:20:18
Job time : 14.6667 secs

```
Db          1 DFT 3


RESULT 11
US-09-428-082B-271
; Sequence 271, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 271
;   LENGTH: 5
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: IL-1 ANTAGONIST PEPTIDE
US-09-428-082B-271

  Query Match           48.6%;  Score 17;  DB 4;  Length 5;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TDF 3
              |||
Db          2 TDF 4


RESULT 12
US-09-233-857-7
; Sequence 7, Application US/09233857
; Patent No. 6495353
; GENERAL INFORMATION:
;  APPLICANT: Plowman, Gregory
;  APPLICANT: Flanagan, Peter
;  TITLE OF INVENTION: HUMAN ORTHOLOGUES OF WART
;  FILE REFERENCE: 239/251
;  CURRENT APPLICATION NUMBER: US/09/233,857
;  CURRENT FILING DATE: 1999-01-20
;  EARLIER APPLICATION NUMBER: USSN 60/072,023
;  EARLIER FILING DATE: 1998-01-21
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: HUMAN
US-09-233-857-7

  Query Match           48.6%;  Score 17;  DB 4;  Length 6;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TDF 3
              |||
Db          3 TDF 5


RESULT 13
US-08-136-743B-29
; Sequence 29, Application US/08136743B
; Patent No. 5459063
;  GENERAL INFORMATION:
;     APPLICANT:  Barry S. Cooperman, Harvey Rubin,
;     APPLICANT:  Jerome Salem, and Alison L. Fisher
;     TITLE OF INVENTION:  "Plasmodium falciparum Ribonu-
;     TITLE OF INVENTION:  cleotide Reductase, DNA Sequences Therefor and Peptide i
;     TITLE OF INVENTION:  Thereof"
;     NUMBER OF SEQUENCES:  67
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  The University of Pennsylvania
;       STREET:  Suite 330
;       STREET:  3700 Market Street
;       CITY:  Philadelphia
;       STATE:  Pennsylvania
;       COUNTRY:  U.S.A.
;       ZIP:  19104-3246
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/136,743B
;       FILING DATE:  10/14/93
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Monaco, Daniel A.
;       REGISTRATION NUMBER:  30,480
;       REFERENCE/DOCKET NUMBER:  3957-10
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (215) 568-8383
;       TELEFAX:  (215) 568-5549
;       TELEX:  No. 5459063e
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  7 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-136-743B-29

  Query Match           48.6%;  Score 17;  DB 1;  Length 7;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TDF 3
              |||
Db          5 TDF 7


RESULT 14
US-08-136-743B-30
; Sequence 30, Application US/08136743B
; Patent No. 5459063
;  GENERAL INFORMATION:
;     APPLICANT:  Barry S. Cooperman, Harvey Rubin,
;     APPLICANT:  Jerome Salem, and Alison L. Fisher
;     TITLE OF INVENTION:  "Plasmodium falciparum Ribonu-
;     TITLE OF INVENTION:  cleotide Reductase, DNA Sequences Therefor and Peptide ir
;     TITLE OF INVENTION:  Thereof"
;     NUMBER OF SEQUENCES:  67
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  The University of Pennsylvania
;       STREET:  Suite 330
;       STREET:  3700 Market Street
;       CITY:  Philadelphia
;       STATE:  Pennsylvania
;       COUNTRY:  U.S.A.
;       ZIP:  19104-3246
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/136,743B
```

```
;      REFERENCE/DOCKET NUMBER:  001560-215
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..5
;      OTHER INFORMATION:  /note= "Amino acid sequence
;      OTHER INFORMATION:  encoded by nucleotides 1-15 of SEQ ID NO. 7."
US-08-910-990-10

  Query Match             48.6%;  Score 17;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TDF 3
              |||
Db          2 TDF 4


RESULT 9
US-09-367-777-132
; Sequence 132, Application US/09367777
; Patent No. 6562598
;   GENERAL INFORMATION:
;        APPLICANT: Himmelspach, Michele
;                   Pfleiderer, Michael
;                   Falkner, Falko-Guenter
;                   Eibl, Johann
;                   Dorner, Friedrich
;                   Schlokat, Uwe
;        TITLE OF INVENTION: Factor X Deletion Mutants
;                            and Analogues Thereof
;        NUMBER OF SEQUENCES: 145
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Townsend and Townsend and Crew LLP
;             STREET: Two Embarcadero Center, Eighth Floor
;             CITY: San Francisco
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94111-3834
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/367,777
;             FILING DATE: 10-No. 6562598-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: AT A 336/97
;             FILING DATE: 27-FEB-1997
;             APPLICATION NUMBER: WO PCT/AT98/00046
;             FILING DATE: 27-FEB-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ausenhus, Scott L.
;             REGISTRATION NUMBER: 42,271
;             REFERENCE/DOCKET NUMBER: 20695D-000900US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-576-0200
;             TELEFAX: 415-576-0300
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 132:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 5 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 132:
US-09-367-777-132

  Query Match             48.6%;  Score 17;  DB 4;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DFT 4
              |||
Db          1 DFT 3


RESULT 10
US-09-367-791A-76
; Sequence 76, Application US/09367791A
; Patent No. 6573071
;   GENERAL INFORMATION:
;        APPLICANT: Himmelspach, Michele
;                   Schlokat, Uwe
;                   Dorner, Friedrich
;                   Fisch, Andreas
;                   Eibl, Johann
;        TITLE OF INVENTION: Factor X Analogues With
;                            a Modified Protease Cleavage Site
;        NUMBER OF SEQUENCES: 122
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Townsend and Townsend and Crew LLP
;             STREET: Two Embarcadero Center, Eighth Floor
;             CITY: San Francisco
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94111-3834
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/367,791A
;             FILING DATE: 12-No. 6573071-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: AT A 335/97
;             FILING DATE: 27-FEB-1997
;             APPLICATION NUMBER: WO PCT/AT98/00045
;             FILING DATE: 27-FEB-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ausenhus, Scott L.
;             REGISTRATION NUMBER: 42,471
;             REFERENCE/DOCKET NUMBER: 20695D-000700US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 576-0200
;             TELEFAX: (415) 576-0300
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 76:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 5 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 76:
US-09-367-791A-76

  Query Match             48.6%;  Score 17;  DB 4;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DFT 4
```

```
                  |||::
Db          1 FTLSV 5


RESULT 6
US-09-155-613A-59
; Sequence 59, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
;  APPLICANT: Boulanger, Pierre
;  APPLICANT:  Hong, Saw See
;  APPLICANT:  Karayan, Lucie
;  TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
;  FILE REFERENCE: 032751-036
;  CURRENT APPLICATION NUMBER: US/09/155,613A
;  CURRENT FILING DATE:  1998-09-30
;  PRIOR APPLICATION NUMBER: PCT/FR98/00184
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: FR 97/01005
;  PRIOR FILING DATE: 1997-01-30
;  PRIOR APPLICATION NUMBER: FR 97/11166
;  PRIOR FILING DATE: 1997-09-09
;  NUMBER OF SEQ ID NOS: 98
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Phagotope
US-09-155-613A-59

  Query Match             51.4%;  Score 18;  DB 4;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DFTLT 6
                ::|||
Db          2 NYTLT 6


RESULT 7
US-08-405-230-10
; Sequence 10, Application US/08405230
; Patent No. 5707846
;  GENERAL INFORMATION:
;    APPLICANT:  TANIGUCHI, Naoyuki
;    APPLICANT:  NISHIKAWA, Atsushi
;    APPLICANT:  YAMAGUCHI, No. 5707846omi
;    TITLE OF INVENTION:  NOVEL N-ACETYLGLUCOSAMINYL TRANSFERASE
;    TITLE OF INVENTION:  GENE CODING THEREFOR AND PROCESS FOR PRODUCTION THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker & Mathis
;      STREET:  George Mason Bldg., Washington & Prince Sts.
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/405,230
;      FILING DATE:  16-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/110,736
;      FILING DATE:  23-AUG-1993
;      APPLICATION NUMBER:  JP 4-245950
;      FILING DATE:  24-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-237118
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Crane-Feury, Sharon E
;      REGISTRATION NUMBER:  36,113
;      REFERENCE/DOCKET NUMBER:  001560-215
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..5
;      OTHER INFORMATION:  /note= "Amino acid sequence
;      OTHER INFORMATION:  encoded by nucleotides 1-15 of SEQ ID NO. 7."
US-08-405-230-10

  Query Match             48.6%;  Score 17;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TDF 3
                |||
Db          2 TDF 4


RESULT 8
US-08-910-990-10
; Sequence 10, Application US/08910990
; Patent No. 5834284
;  GENERAL INFORMATION:
;    APPLICANT:  TANIGUCHI, Naoyuki
;    APPLICANT:  NISHIKAWA, Atsushi
;    APPLICANT:  YAMAGUCHI, No. 5834284omi
;    TITLE OF INVENTION:  NOVEL N-ACETYLGLUCOSAMINYL TRANSFERASE
;    TITLE OF INVENTION:  GENE CODING THEREFOR AND PROCESS FOR PRODUCTION THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker & Mathis
;      STREET:  George Mason Bldg., Washington & Prince Sts.
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/910,990
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/405,230
;      FILING DATE:  16-MAR-1995
;      APPLICATION NUMBER:  US 08/110,736
;      FILING DATE:  23-AUG-1993
;      APPLICATION NUMBER:  JP 4-245950
;      FILING DATE:  24-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-237118
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Crane-Feury, Sharon E
;      REGISTRATION NUMBER:  36,113
```

```
;   LOCATION: ()..()
;   OTHER INFORMATION: Epitope tag
US-09-724-297-4

  Query Match             54.3%;  Score 19;  DB 4;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 3e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFTL 5
              ||| |
Db          1 TDFYL 5


RESULT 3
US-08-916-443A-8
; Sequence 8, Application US/08916443A
; Patent No. 6001986
;  GENERAL INFORMATION:
;    APPLICANT:  Yong Sig KIM
;    APPLICANT:  Sun Chung PARK
;    APPLICANT:  Soo Kyung OH
;    APPLICANT:  Hosull LEE
;    APPLICANT:  Jeong Woo CHO
;    APPLICANT:  Chang H. CHUNG
;    TITLE OF INVENTION:  Antiviral Proteins, Amarandin 1 and 2, from
;    TITLE OF INVENTION:  Amaranthus Viridis, DNAs Encoding Therefrom
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3+ Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  WordPerfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/916,443A
;      FILING DATE:  22 AUG 1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  1942/18
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-429-1776
;      TELEFAX:  202-429-0796
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-916-443A-8

  Query Match             54.3%;  Score 19;  DB 3;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 DFTLTI 7
              | | |:
Db          2 DLTFTV 7


RESULT 4
US-08-640-737-38
; Sequence 38, Application US/08640737
; Patent No. 6215044
;  GENERAL INFORMATION:
;    APPLICANT:  ARROWSMITH, David A.
;    APPLICANT:  HELLYER, Susan A.
;    APPLICANT:  DE SILVA, Jacqueline
;    APPLICANT:  WHITEMAN, Sally A.
;    TITLE OF INVENTION:  Tomato Xyloglucan Endo-Transglycosylase
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pillsbury Madison & Sutro, L.L.P.
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MS Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/640,737
;      FILING DATE:  06-MAY-1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB94/02467
;      FILING DATE:  10-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9323225.4
;      FILING DATE:  10-NOV-1993
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
US-08-640-737-38

  Query Match             54.3%;  Score 19;  DB 3;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 3e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DFTLT 6
              | |||
Db          3 DHTLT 7


RESULT 5
US-08-753-750B-20
; Sequence 20, Application US/08753750B
; Patent No. 6610506
; GENERAL INFORMATION:
;  APPLICANT: Lo, Reggie Y.C.
;  APPLICANT:  Schryvers, Anthony B.
;  APPLICANT:  Potter, Andrew A.
;  TITLE OF INVENTION: TRANSFERRIN BINDING PROTEINS OF
;  TITLE OF INVENTION:  PASTEURELLA HAEMOLYTICA AND VACCINES CONTAINING THE SAME
;  FILE REFERENCE: A34762 021645.0105
;  CURRENT APPLICATION NUMBER: US/08/753,750B
;  CURRENT FILING DATE:  1996-11-29
;  PRIOR APPLICATION NUMBER: CA 2,164,274
;  PRIOR FILING DATE: 1995-12-01
;  PRIOR APPLICATION NUMBER: 60/008,569
;  PRIOR FILING DATE: 1995-12-01
;  NUMBER OF SEQ ID NOS: 68
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-08-753-750B-20

  Query Match             54.3%;  Score 19;  DB 4;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FTLTI 7
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 11:13:51 ; Search time 13.6667 Seconds
(without alignments)
26.443 Million cell updates/sec

Title: US-09-712-819C-5
Perfect score: 35
Sequence: 1 TDFTLTI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 57228

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	19	54.3	6	4	US-09-140-084-4	Sequence 4, Appli
2	19	54.3	6	4	US-09-724-297-4	Sequence 4, Appli
3	19	54.3	7	3	US-08-916-443A-8	Sequence 8, Appli
4	19	54.3	7	3	US-08-640-737-38	Sequence 38, Appl
5	19	54.3	7	4	US-08-753-750B-20	Sequence 20, Appl
6	18	51.4	6	4	US-09-155-613A-59	Sequence 59, Appl
7	17	48.6	5	1	US-08-405-230-10	Sequence 10, Appl
8	17	48.6	5	2	US-08-910-990-10	Sequence 10, Appl
9	17	48.6	5	4	US-09-367-777-132	Sequence 132, App
10	17	48.6	5	4	US-09-367-791A-76	Sequence 76, Appl
11	17	48.6	5	4	US-09-428-082B-271	Sequence 271, App
12	17	48.6	6	4	US-09-233-857-7	Sequence 7, Appli
13	17	48.6	7	1	US-08-136-743B-29	Sequence 29, Appl
14	17	48.6	7	1	US-08-136-743B-30	Sequence 30, Appl
15	17	48.6	7	1	US-08-136-743B-31	Sequence 31, Appl
16	17	48.6	7	1	US-08-136-743B-32	Sequence 32, Appl
17	17	48.6	7	1	US-08-405-230-5	Sequence 5, Appli
18	17	48.6	7	2	US-08-910-990-5	Sequence 5, Appli
19	17	48.6	7	2	US-08-739-401A-6	Sequence 6, Appli
20	17	48.6	7	5	PCT-US93-11703-72	Sequence 72, Appl
21	16	45.7	5	1	US-08-180-209B-14	Sequence 14, Appl
22	16	45.7	5	1	US-08-385-745-14	Sequence 14, Appl
23	16	45.7	5	3	US-08-591-632-17	Sequence 17, Appl
24	16	45.7	5	3	US-08-591-632-23	Sequence 23, Appl
25	16	45.7	5	3	US-08-591-632-26	Sequence 26, Appl
26	16	45.7	5	3	US-08-485-388-14	Sequence 14, Appl
27	16	45.7	5	3	US-08-474-853-14	Sequence 14, Appl
28	16	45.7	5	4	US-09-166-205B-14	Sequence 14, Appl
29	16	45.7	5	4	US-09-611-451-17	Sequence 17, Appl
30	16	45.7	5	4	US-09-611-451-23	Sequence 23, Appl
31	16	45.7	5	4	US-09-611-451-26	Sequence 26, Appl
32	16	45.7	5	5	PCT-US94-02629-14	Sequence 14, Appl
33	16	45.7	6	1	US-08-252-995D-7	Sequence 7, Appli
34	16	45.7	6	2	US-08-482-228-180	Sequence 180, App
35	16	45.7	6	2	US-08-834-108-7	Sequence 7, Appli
36	16	45.7	6	3	US-08-482-528-180	Sequence 180, App
37	16	45.7	6	4	US-09-535-852-1745	Sequence 1745, Ap
38	16	45.7	7	1	US-08-136-743B-55	Sequence 55, Appl
39	16	45.7	7	1	US-08-096-946-5	Sequence 5, Appli
40	16	45.7	7	2	US-08-177-109A-7	Sequence 7, Appli
41	16	45.7	7	2	US-08-687-706-7	Sequence 7, Appli
42	16	45.7	7	3	US-09-040-216-28	Sequence 28, Appl
43	16	45.7	7	3	US-09-173-941-52	Sequence 52, Appl
44	16	45.7	7	4	US-09-494-190-52	Sequence 52, Appl
45	16	45.7	7	4	US-09-535-852-1746	Sequence 1746, Ap

ALIGNMENTS

RESULT 1
US-09-140-084-4
; Sequence 4, Application US/09140084A
; Patent No. 6300065
; GENERAL INFORMATION:
; APPLICANT: Kieke, et al.
; TITLE OF INVENTION: Yeast Cell Surface Display of Proteins and Uses Thereof
; FILE REFERENCE: D6061CIP2
; CURRENT APPLICATION NUMBER: US/09/140,084A
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:Epitope Tag
US-09-140-084-4

Query Match 54.3%; Score 19; DB 4; Length 6;
Best Local Similarity 80.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TDFTL 5
              ||| |
Db          1 TDFYL 5
```

RESULT 2
US-09-724-297-4
; Sequence 4, Application US/09724297
; Patent No. 6423538
; GENERAL INFORMATION:
; APPLICANT: The Board of Trustees of the University of Illinois
; APPLICANT: Wittrup, et al.
; TITLE OF INVENTION: Yeast Cell Surface Display of Proteins and Uses Thereof
; FILE REFERENCE: 97-99C
; CURRENT APPLICATION NUMBER: US/09/724,297
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/009,388
; PRIOR FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 6
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: misc_feature

```
  Query Match             51.4%;  Score 18;  DB 14;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TLTI 7
              ||||
Db          1 TLTI 4


RESULT 14
US-09-788-006-107
; Sequence 107, Application US/09788006
; Publication No. US20030036093A1
; GENERAL INFORMATION:
;  APPLICANT: Floudas, Christopher A.
;  APPLICANT:  Klepeis, John L.
;  TITLE OF INVENTION: Methods of Ab Initio Prediction of Alpha Helices, Beta Sheets,and
;  TITLE OF INVENTION:  Polypeptide Tertiary Structures
;  FILE REFERENCE: PU-0007
;  CURRENT APPLICATION NUMBER: US/09/788,006
;  CURRENT FILING DATE:  2001-02-16
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 107
;   LENGTH: 5
;   TYPE: PRT
;   ORGANISM: Streptomyces griseus
US-09-788-006-107

  Query Match             48.6%;  Score 17;  DB 10;  Length 5;
  Best Local Similarity   75.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FTLT 6
              ||:|
Db          2 FTVT 5


RESULT 15
US-09-788-006-108
; Sequence 108, Application US/09788006
; Publication No. US20030036093A1
; GENERAL INFORMATION:
;  APPLICANT: Floudas, Christopher A.
;  APPLICANT:  Klepeis, John L.
;  TITLE OF INVENTION: Methods of Ab Initio Prediction of Alpha Helices, Beta Sheets,and
;  TITLE OF INVENTION:  Polypeptide Tertiary Structures
;  FILE REFERENCE: PU-0007
;  CURRENT APPLICATION NUMBER: US/09/788,006
;  CURRENT FILING DATE:  2001-02-16
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
;   LENGTH: 5
;   TYPE: PRT
;   ORGANISM: Streptomyces griseus
US-09-788-006-108

  Query Match             48.6%;  Score 17;  DB 10;  Length 5;
  Best Local Similarity   75.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FTLT 6
              ||:|
Db          1 FTVT 4


Search completed: June  1, 2004, 11:31:23
Job time : 35.6667 secs
```

; APPLICANT: Mazur, Wieslaw A
; TITLE OF INVENTION: Cortcotropin Releasing Factor 2 Receptor Agonists
; FILE REFERENCE: 8847M
; CURRENT APPLICATION NUMBER: US/10/317,252A
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/349,117
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/376,337
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/388,895
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/411,988
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 530
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 402
; LENGTH: 6
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-317-252A-402

Query Match 54.3%; Score 19; DB 14; Length 6;
Best Local Similarity 80.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TDFTL 5
| |||
Db 2 TKFTL 6

RESULT 11
US-10-156-820-59
; Sequence 59, Application US/10156820
; Publication No. US20020150558A1
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/10/156,820
; CURRENT FILING DATE: 2002-06-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotope
US-10-156-820-59

Query Match 51.4%; Score 18; DB 13; Length 6;
Best Local Similarity 60.0%; Pred. No. 1e+06;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DFTLT 6
::|||
Db 2 NYTLT 6

RESULT 12
US-10-267-565-11
; Sequence 11, Application US/10267565
; Publication No. US20030204059A1
; GENERAL INFORMATION:
; APPLICANT: Gately, Maurice
; APPLICANT: Gubler, Ulrich
; APPLICANT: Hulmes, Jeffery
; APPLICANT: Podlaski, Frank
; APPLICANT: Stern, Alvin
; TITLE OF INVENTION: PURIFICATION AND CHARACTERIZATION OF CYTOTOXIC
; TITLE OF INVENTION: LYMPHOCYTE MATURATION FACTOR AND MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES THERETO
; FILE REFERENCE: 11126-005
; CURRENT APPLICATION NUMBER: US/10/267,565
; CURRENT FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: 09/401,839
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: 08/459,151
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: 08/205,011
; PRIOR FILING DATE: 1994-03-02
; PRIOR APPLICATION NUMBER: 07/857,023
; PRIOR FILING DATE: 1992-03-24
; PRIOR APPLICATION NUMBER: 07/572,284
; PRIOR FILING DATE: 1990-08-27
; PRIOR APPLICATION NUMBER: 07/520,935
; PRIOR FILING DATE: 1990-05-09
; PRIOR APPLICATION NUMBER: 07/455,708
; PRIOR FILING DATE: 1989-12-22
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-565-11

Query Match 51.4%; Score 18; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TLTI 7
||||
Db 1 TLTI 4

RESULT 13
US-10-349-507-11
; Sequence 11, Application US/10349507
; Publication No. US20030199002A1
; GENERAL INFORMATION:
; APPLICANT: Hekimi , Siegfried
; APPLICANT: Jiang , Ning
; APPLICANT: Benard, Claire
; APPLICANT: Kebir , Hania
; APPLICANT: McCright, Brenton
; APPLICANT: Lakowski , Bernard
; TITLE OF INVENTION: CLK-2 NUCLEIC ACIDS, POLYPEPTIDES, AND USES THEREOF
; FILE REFERENCE: 11202-006-999
; CURRENT APPLICATION NUMBER: US/10/349,507
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 10/312,187
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/CA01/00913
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/213,174
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/254,932
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-349-507-11

```
 ; Publication No. US20040072739A1
 ; GENERAL INFORMATION:
 ;  APPLICANT: Anderson, Christen M.
 ;  APPLICANT:  Clevenger, William
 ;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING
 ;  TITLE OF INVENTION:  ENDOGENOUS INHIBITOR OF ATP SYNTHASE, INCLUDING
 ;  TITLE OF INVENTION:  TREATMENT FOR DIABETES
 ;  FILE REFERENCE: 660088.435C1
 ;  CURRENT APPLICATION NUMBER: US/09/796,076
 ;  CURRENT FILING DATE:  2001-02-27
 ;  NUMBER OF SEQ ID NOS: 68
 ;  SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 4
 ;   LENGTH: 6
 ;   TYPE: PRT
 ;   ORGANISM: Artificial Sequence
 ;   FEATURE:
 ;   OTHER INFORMATION: Epitope tag
US-09-796-076-4

  Query Match             54.3%;  Score 19;  DB 12;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFTL 5
              ||| |
Db          1 TDFYL 5


RESULT 7
US-10-083-815-4
; Sequence 4, Application US/10083815
; Publication No. US20030026781A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, Christen M.
;  APPLICANT:  Clevenger, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING
;  TITLE OF INVENTION:  ENDOGENOUS INHIBITOR OF ATP SYNTHASE, INCLUDING
;  TITLE OF INVENTION:  TREATMENT FOR DIABETES
;  FILE REFERENCE: 660088.435C2
;  CURRENT APPLICATION NUMBER: US/10/083,815
;  CURRENT FILING DATE:  2002-02-27
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Epitope tag
US-10-083-815-4

  Query Match             54.3%;  Score 19;  DB 14;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFTL 5
              ||| |
Db          1 TDFYL 5


RESULT 8
US-10-315-964A-402
; Sequence 402, Application US/10315964A
; Publication No. US20030148956A1
; GENERAL INFORMATION:
;  APPLICANT: The Procter & Gamble Company
;  APPLICANT:  Isfort, Robert J
;  APPLICANT:  Mazur, Wieslaw A
;  TITLE OF INVENTION: Cortcotropin Releasing Factor 2 Receptor Agonists
;  FILE REFERENCE: 8847M3
;  CURRENT APPLICATION NUMBER: US/10/315,964A
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/349,117
;  PRIOR FILING DATE: 2002-01-16
;  PRIOR APPLICATION NUMBER: US 60/376,337
;  PRIOR FILING DATE: 2002-04-29
;  PRIOR APPLICATION NUMBER: US 60/388,895
;  PRIOR FILING DATE: 2002-06-14
;  PRIOR APPLICATION NUMBER: US 60/411,988
;  PRIOR FILING DATE: 2002-09-19
;  NUMBER OF SEQ ID NOS: 530
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 402
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: artificial
;   FEATURE:
;   OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-315-964A-402

  Query Match             54.3%;  Score 19;  DB 14;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFTL 5
              | |||
Db          2 TKFTL 6


RESULT 9
US-10-317-251A-402
; Sequence 402, Application US/10317251A
; Publication No. US20030148957A1
; GENERAL INFORMATION:
;  APPLICANT: The Procter & Gamble Company
;  APPLICANT:  Isfort, Robert J
;  APPLICANT:  Mazur, Wieslaw A
;  TITLE OF INVENTION: Cortcotropin Releasing Factor 2 Receptor Agonists
;  FILE REFERENCE: 8847M2
;  CURRENT APPLICATION NUMBER: US/10/317,251A
;  CURRENT FILING DATE:  2002-12-11
;  PRIOR APPLICATION NUMBER: US 60/349,117
;  PRIOR FILING DATE: 2002-01-16
;  PRIOR APPLICATION NUMBER: US 60/376,337
;  PRIOR FILING DATE: 2002-04-29
;  PRIOR APPLICATION NUMBER: US 60/388,895
;  PRIOR FILING DATE: 2002-06-14
;  PRIOR APPLICATION NUMBER: US 60/411,988
;  PRIOR FILING DATE: 2002-09-19
;  NUMBER OF SEQ ID NOS: 530
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 402
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: artificial
;   FEATURE:
;   OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-317-251A-402

  Query Match             54.3%;  Score 19;  DB 14;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFTL 5
              | |||
Db          2 TKFTL 6


RESULT 10
US-10-317-252A-402
; Sequence 402, Application US/10317252A
; Publication No. US20030148958A1
; GENERAL INFORMATION:
;  APPLICANT: The Procter & Gamble Company
;  APPLICANT:  Isfort, Robert J
```

```
; APPLICANT: Mazur, Wieslaw A
; TITLE OF INVENTION: Cortcotropin Releasing Factor 2 Receptor Agonists
; FILE REFERENCE: 8847M3
; CURRENT APPLICATION NUMBER: US/10/315,964A
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/349,117
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/376,337
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/388,895
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/411,988
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 530
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 394
; LENGTH: 5
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-315-964A-394

  Query Match             54.3%;  Score 19;  DB 14;  Length 5;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFTL 5
              | |||
Db          1 TKFTL 5


RESULT 3
US-10-317-251A-394
; Sequence 394, Application US/10317251A
; Publication No. US20030148957A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Isfort, Robert J
; APPLICANT: Mazur, Wieslaw A
; TITLE OF INVENTION: Cortcotropin Releasing Factor 2 Receptor Agonists
; FILE REFERENCE: 8847M2
; CURRENT APPLICATION NUMBER: US/10/317,251A
; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/349,117
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/376,337
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/388,895
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/411,988
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 530
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 394
; LENGTH: 5
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-317-251A-394

  Query Match             54.3%;  Score 19;  DB 14;  Length 5;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFTL 5
              | |||
Db          1 TKFTL 5


RESULT 4
US-10-317-252A-394
; Sequence 394, Application US/10317252A
; Publication No. US20030148958A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Isfort, Robert J
; APPLICANT: Mazur, Wieslaw A
; TITLE OF INVENTION: Cortcotropin Releasing Factor 2 Receptor Agonists
; FILE REFERENCE: 8847M
; CURRENT APPLICATION NUMBER: US/10/317,252A
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/349,117
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/376,337
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/388,895
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/411,988
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 530
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 394
; LENGTH: 5
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-317-252A-394

  Query Match             54.3%;  Score 19;  DB 14;  Length 5;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFTL 5
              | |||
Db          1 TKFTL 5


RESULT 5
US-10-351-891-122
; Sequence 122, Application US/10351891
; Publication No. US20040048311A1
; GENERAL INFORMATION:
; APPLICANT: DANA AULT-RICHE
; APPLICANT: PAUL D. KASSNER
; TITLE OF INVENTION: USE OF COLLECTIONS OF BINDING SITES FOR SAMPLE PROFILING AI
; TITLE OF INVENTION: APPLICATIONS
; FILE REFERENCE: 25885-1753
; CURRENT APPLICATION NUMBER: US/10/351,891
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/352,011
; PRIOR FILING DATE: 2002-01-24
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: AU5 Peptide
US-10-351-891-122

  Query Match             54.3%;  Score 19;  DB 12;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFTL 5
              ||| |
Db          1 TDFYL 5


RESULT 6
US-09-796-076-4
; Sequence 4, Application US/09796076
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 11:18:42 ; Search time 34.6667 Seconds
(without alignments)
56.387 Million cell updates/sec

Title: US-09-712-819C-5
Perfect score: 35
Sequence: 1 TDFTLTI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 49082

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	57.1	6	11	US-09-943-944E-126	Sequence 126, App
2	19	54.3	5	14	US-10-315-964A-394	Sequence 394, App
3	19	54.3	5	14	US-10-317-251A-394	Sequence 394, App
4	19	54.3	5	14	US-10-317-252A-394	Sequence 394, App
5	19	54.3	6	12	US-10-351-891-122	Sequence 122, App
6	19	54.3	6	12	US-09-796-076-4	Sequence 4, Appli
7	19	54.3	6	14	US-10-083-815-4	Sequence 4, Appli
8	19	54.3	6	14	US-10-315-964A-402	Sequence 402, App
9	19	54.3	6	14	US-10-317-251A-402	Sequence 402, App
10	19	54.3	6	14	US-10-317-252A-402	Sequence 402, App
11	18	51.4	6	13	US-10-156-820-59	Sequence 59, Appl
12	18	51.4	7	12	US-10-267-565-11	Sequence 11, Appl
13	18	51.4	7	14	US-10-349-507-11	Sequence 11, Appl
14	17	48.6	5	10	US-09-788-006-107	Sequence 107, App
15	17	48.6	5	10	US-09-788-006-108	Sequence 108, App
16	17	48.6	5	12	US-10-436-549-151	Sequence 151, App
17	17	48.6	5	12	US-10-436-549-200	Sequence 200, App
18	17	48.6	5	12	US-10-609-217-271	Sequence 271, App
19	17	48.6	5	12	US-10-632-388-271	Sequence 271, App
20	17	48.6	5	12	US-10-651-723-271	Sequence 271, App
21	17	48.6	5	12	US-10-645-761-271	Sequence 271, App
22	17	48.6	5	14	US-10-348-504-132	Sequence 132, App
23	17	48.6	5	14	US-10-407-123-76	Sequence 76, Appl
24	17	48.6	5	16	US-10-666-696-271	Sequence 271, App
25	17	48.6	5	16	US-10-653-048-271	Sequence 271, App
26	17	48.6	6	9	US-09-770-102A-55	Sequence 55, Appl
27	17	48.6	6	12	US-10-284-130-7	Sequence 7, Appli
28	17	48.6	6	14	US-10-172-919-25	Sequence 25, Appl
29	17	48.6	7	10	US-09-954-385-257	Sequence 257, App
30	17	48.6	7	16	US-10-264-309-288	Sequence 288, App
31	16	45.7	5	12	US-10-380-533-126	Sequence 126, App
32	16	45.7	6	14	US-10-286-993-5	Sequence 5, Appli
33	16	45.7	7	9	US-09-832-723-69	Sequence 69, Appl
34	16	45.7	7	13	US-10-080-100-93	Sequence 93, Appl
35	16	45.7	7	14	US-10-303-331-69	Sequence 69, Appl
36	15	42.9	5	10	US-09-788-006-2	Sequence 2, Appli
37	15	42.9	5	10	US-09-788-006-3	Sequence 3, Appli
38	15	42.9	5	14	US-10-214-796-21	Sequence 21, Appl
39	15	42.9	5	14	US-10-286-186-3	Sequence 3, Appli
40	15	42.9	5	14	US-10-286-186-4	Sequence 4, Appli
41	15	42.9	6	9	US-09-293-854-8	Sequence 8, Appli
42	15	42.9	6	10	US-09-990-586-8	Sequence 8, Appli
43	15	42.9	6	12	US-10-400-991-75	Sequence 75, Appl
44	15	42.9	6	14	US-10-105-930-39	Sequence 39, Appl
45	15	42.9	6	14	US-10-006-869-650	Sequence 650, App

ALIGNMENTS

RESULT 1
US-09-943-944E-126
; Sequence 126, Application US/09943944E
; Publication No. US20040014036A1
; GENERAL INFORMATION:
; APPLICANT: Ptashne, et al.,
; TITLE OF INVENTION: Transcriptional Activation System, Activators, and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: 0342941-0065
; CURRENT APPLICATION NUMBER: US/09/943,944E
; CURRENT FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 126
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Random peptide
; OTHER INFORMATION: sequences.
US-09-943-944E-126

Query Match 57.1%; Score 20; DB 11; Length 6;
Best Local Similarity 80.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TDFTL 5
              ||| |
Db          1 TDFLL 5
```

RESULT 2
US-10-315-964A-394
; Sequence 394, Application US/10315964A
; Publication No. US20030148956A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Isfort, Robert J

XX
SQ Sequence 6 AA;

Query Match 54.3%; Score 19; DB 3; Length 6;
Best Local Similarity 80.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TDFTL 5
              |||| |
Db          1 TDFYL 5
```

Search completed: June 1, 2004, 11:16:16
Job time : 47.6667 secs

XX
PR 20-JAN-1998; 98US-00009388.
PR 26-AUG-1998; 98US-00140084.
XX
PA (UNII) UNIV ILLINOIS FOUND.
XX
PI Wittrup KD, Kieke MC, Kranz DM, Shusta E, Boder ET;
XX
DR WPI; 1999-430619/36.
XX
PT Selecting proteins with enhanced phenotypic properties than wild-type
PT proteins, is useful for highly specific cancer diagnosis and therapy.
XX
PS Disclosure; Page 7; 116pp; English.
XX
CC This peptide comprises an epitope tag that can be used in methods of the
CC invention. The invention discloses a powerful new system for engineering
CC antibody affinity and specificity, by constructing a microbial analogue
CC of the mammalian system's B cell repertoire. Antibodies are displayed on
CC the surface of yeast cells by genetic fusion with yeast cell wall
CC proteins, especially agglutinin proteins. After mutation, variants are
CC selected on the basis of improved binding characteristics with
CC fluorescently labeled targets. The selection method also identifies
CC proteins with enhanced phenotypic characteristics, proteins that are
CC displayed at higher levels, proteins that are secreted at higher
CC efficiency and proteins of improved stability
XX
SQ Sequence 6 AA;

Query Match 54.3%; Score 19; DB 2; Length 6;
Best Local Similarity 80.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TDFTL 5
||| |
Db 1 TDFYL 5

RESULT 14
AAE16558
ID AAE16558 standard; peptide; 6 AA.
XX
AC AAE16558;
XX
DT 09-APR-2002 (first entry)
XX
DE Epitope tag #2 fused to yeast cell wall protein AGA2.
XX
KW Yeast cell wall protein; AGA2; T cell receptor; multiple sclerosis;
KW cancer; sepsis; autoimmune disease; arthritis; diabetes.
XX
OS Synthetic.
XX
PN US6331391-B1.
XX
PD 18-DEC-2001.
XX
PF 20-JAN-1998; 98US-00009388.
XX
PR 31-MAY-1996; 96US-0018741P.
PR 30-MAY-1997; 97US-00866398.
XX
PA (UNII) UNIV ILLINOIS FOUND.
XX
PI Wittrup KD, Kranz DM, Kieke M, Boder ET;
XX
DR WPI; 1999-430619/36.
XX
PT Selecting proteins with enhanced phenotypic properties than wild-type
PT proteins, is useful for highly specific cancer diagnosis and therapy.
XX
PS Claim 40; Col 60; 59pp; English.
XX
CC The present invention relates to a method for selecting proteins for
CC displayability on a yeast cell surface. The method comprises transforming
CC yeast cells with a vector that expresses a test protein fused to a yeast
CC cell wall protein (AGA2), contacting the cells with a label that binds to
CC proteins displayed on the cell wall, and isolating label-bound cells,
CC where the test protein is from a variegated population generated by
CC mutagenesis. The invention is also directed to new processes for
CC engineering T cell receptor for improved binding properties. Improved T
CC cell receptor molecules are useful in therapies for cancer, sepsis,
CC autoimmune diseases such as arthritis, diabetes or multiple sclerosis.
CC The methods are useful to select proteins with altered affinity, altered
CC specificity or conditional binding. The present sequence is an epitope
CC tag fused between protein of interest and yeast cell wall protein AGA2
XX
SQ Sequence 6 AA;

Query Match 54.3%; Score 19; DB 2; Length 6;
Best Local Similarity 80.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TDFTL 5
||| |
Db 1 TDFYL 5

RESULT 15
AAY77708
ID AAY77708 standard; peptide; 6 AA.
XX
AC AAY77708;
XX
DT 12-MAY-2000 (first entry)
XX
DE AU5 peptide epitope.
XX
KW Cell surface receptor; luminescence; protein internalization;
KW drug discovery; screening assay; epitope; AU5.
XX
OS Synthetic.
XX
PN WO200003246-A2.
XX
PD 20-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015870.
XX
PR 13-JUL-1998; 98US-0092671P.
XX
PA (CELL-) CELLOMICS INC.
XX
PI Rubin RA, Giuliano KA, Gough A, Dunlay T;
XX
DR WPI; 2000-171170/15.
XX
PT Automated screening method for identifying compounds which induce cell
PT surface receptor internalization, useful for drug discovery.
XX
PS Example 6; Page 67; 148pp; English.
XX
CC The invention relates to a method for identifying compounds which inhibit
CC internalization of cell surface receptors. Provided are an array of
CC locations, each containing cells with a cell surface receptor protein,
CC that are treated with a test compound. The protein is luminescently
CC labeled or contacted with a luminescently labeled cell before or after
CC test compound treatment. Any luminescence produced is converted into
CC digital data and automatically analysed to determine if the test compound
CC induced the protein internalization. The novel method is used to screen
CC for compounds which modulate cell surface receptor protein
CC internalization, this can be used in drug discovery, to test compound
CC efficacy in living biological systems. The assay method is automated and
CC compact. It has high throughput and uses smaller volumes of reagents and
CC test compounds. Sequences AAY77704-718 represent examples of peptide
CC epitope tags used in the course of the invention

XX
PR 16-JAN-2002; 2002US-0349117P.
PR 29-APR-2002; 2002US-0376337P.
PR 14-JUN-2002; 2002US-0388895P.
PR 19-SEP-2002; 2002US-0411988P.
XX
PA (PROC) PROCTER & GAMBLE CO.
XX
PI Isfort RJ, Mazur WA;
XX
DR WPI; 2003-787975/74.
XX
PT New non-native peptide derived from corticotropin-releasing factor-2,
PT useful for treatment and prevention of e.g. muscular atrophy, also
PT related nucleic acid and antibodies.
XX
PS Example 2; SEQ ID NO 394; 304pp; English.
XX
CC The invention relates to a novel non-native peptide derived from
CC corticotropin-releasing factor-2 (CRF2). The CRF2 peptides have the
CC following activities: myopathic, osteopathic, hypotensive, cardiant,
CC vasotropic, antimigraine, cerebroprotective, nootropic, neuroprotective,
CC anorectic, antidiabetic, analgesic, antiallergic, tranquilizer,
CC anxiolytic, antidepressant, and antiarthritic. The CRF2 peptides, and
CC related compounds derived from other proteins, are used to prevent or
CC treat disorders modulated by the CRF2 receptor, e.g. skeletal muscle
CC atrophy or wasting, and bone disorders, however caused; heart/circulatory
CC diseases (e.g. hypertension, congestive heart failure, heart attack,
CC reperfusion injury, migraine, stroke, memory loss, Alzheimer's disease,
CC dementia); joint disorders (osteoarthritis or rheumatoid arthritis);
CC metabolic disease (obesity or diabetes); pain; allergy; stress; anxiety;
CC low levels of adrenocorticotrophic hormone; anorexia nervosa; depression;
CC also to reduce body temperature and to control appetite or cognitive
CC function. Nucleic acids, optionally labelled, that encode the CRF2
CC peptides are used as primers and probes for amplification, also for gene
CC synthesis and for recombinant production of CRF2 peptides, including use
CC in gene therapy. Antibodies specific for the CRF2 peptides are used to
CC evaluate expression of the CRF2 peptides after gene therapy. This
CC sequence represents a novel native CRF polypeptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 54.3%; Score 19; DB 7; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TDFTL 5
| |||
Db 1 TKFTL 5

RESULT 12
ADE51377
ID ADE51377 standard; peptide; 5 AA.
XX
AC ADE51377;
XX
DT 29-JAN-2004 (first entry)
XX
DE CRF2 non-native polypeptide, SEQ ID No 394.
XX
KW non-native; corticotropin-releasing factor-2; CRF2; myopathic;
KW osteopathic; hypotensive; cardiant; vasotropic; antimigraine;
KW cerebroprotective; nootropic; neuroprotective; anorectic; antidiabetic;
KW analgesic; antiallergic; tranquilizer; anxiolytic; antidepressant;
KW antiarthritic.
XX
OS Unidentified.
XX
PN WO2003062268-A2.
XX
PD 31-JUL-2003.
XX
PF 16-JAN-2003; 2003WO-US001451.
XX
PR 16-JAN-2002; 2002US-0349117P.
PR 29-APR-2002; 2002US-0376337P.
PR 14-JUN-2002; 2002US-0388895P.
PR 19-SEP-2002; 2002US-0411988P.
XX
PA (PROC) PROCTER & GAMBLE CO.
XX
PI Isfort RJ, Mazur WA;
XX
DR WPI; 2003-787974/74.
XX
PT New non-native peptide derived from corticotropin-releasing factor-2,
PT useful for treatment and prevention of e.g. muscular atrophy, also
PT related nucleic acid and antibodies.
XX
PS Example 2; SEQ ID NO 394; 300pp; English.
XX
CC The invention relates to a novel non-native peptide derived from
CC corticotropin-releasing factor-2 (CRF2). The non-native CRF2 peptides
CC have the following activities: myopathic, osteopathic, hypotensive,
CC cardiant, vasotropic, antimigraine, cerebroprotective, nootropic,
CC neuroprotective, anorectic, antidiabetic, analgesic, antiallergic,
CC tranquilizer, anxiolytic, antidepressant, and antiarthritic. The non-
CC native CRF2 peptides, and related compounds derived from other proteins,
CC are used to prevent or treat disorders modulated by the CRF2 receptor,
CC e.g. skeletal muscle atrophy or wasting, and bone disorders, however
CC caused; heart/circulatory diseases (e.g. hypertension, congestive heart
CC failure, heart attack, reperfusion injury, migraine, stroke, memory loss,
CC Alzheimer's diseases, dementia); joint disorders (osteoarthritis or
CC rheumatoid arthritis); metabolic disease (obesity or diabetes); pain;
CC allergy; stress; anxiety; low levels of adrenocorticotrophic hormone;
CC anorexia nervosa; depression; also to reduce body temperature and to
CC control appetite or cognitive function. Nucleic acids, optionally
CC labelled, that encode the non-native CRF2 peptides are used as primers
CC and probes for amplification, also for gene synthesis and for recombinant
CC production of the non-native CRF2 peptides, including use in gene
CC therapy. Antibodies specific for the non-native CRF2 peptides are used to
CC evaluate expression of the non-native CRF2 peptides after gene therapy.
CC This sequence represents a CRF2 non-native polypeptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 54.3%; Score 19; DB 7; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TDFTL 5
| |||
Db 1 TKFTL 5

RESULT 13
AAY06466
ID AAY06466 standard; peptide; 6 AA.
XX
AC AAY06466;
XX
DT 27-SEP-1999 (first entry)
XX
DE Epitope tag.
XX
KW Epitope tag; antibody engineering; yeast; surface display;
KW protein library; peptide library.
XX
OS Synthetic.
XX
PN WO9936569-A1.
XX
PD 22-JUL-1999.
XX
PF 20-JAN-1999; 99WO-US001188.

```
RESULT 9
AAW31467
ID   AAW31467 standard; protein; 6 AA.
XX
AC   AAW31467;
XX
DT   04-AUG-1998  (first entry)
XX
DE   Transcriptional activator peptide fragment LS130.
XX
KW   Activating sequence; Gal4; transcriptional activator; RNA polymerase;
KW   Protein-protein interaction; gene therapy; therapeutic; holoenzyme;
KW   Gal11; DNA binding domain.
XX
OS   Synthetic.
XX
PN   WO9744447-A2.
XX
PD   27-NOV-1997.
XX
PF   02-MAY-1997;   97WO-US007338.
XX
PR   03-MAY-1996;   96US-0017016P.
PR   01-MAY-1997;   97US-00017016.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Ptashne M,  Lu X,  Wu Y;
XX
DR   WPI; 1998-018502/02.
DR   N-PSDB; AAV02565.
XX
PT   New transcriptional activator containing DNA binding domain bound to
PT   peptide - useful for controlling gene expression, especially in gene
PT   therapy, and in protein-protein interaction assays, does not inhibit
PT   other transcription activators.
XX
PS   Example 1; Page 26; 55pp; English.
XX
CC   AAW31412-W31459, AAW31461-W31481, AAW31483-W31493 and AAW39053-W39076 are
CC   fragments used in an assay to determine novel transcriptional activators.
CC   The method involves the production of transcriptional activators
CC   comprising of a DNA-binding group and a 6-25 amino acid peptide that is
CC   covalently bonded to the DNA binding group and does not represent a
CC   fragment of a natural transcription activator. Protein-protein
CC   interactions are identified in the assay by fusing a DNA-binding domain
CC   to a library of DNA fragments and introducing this and a fusion of target
CC   protein and a polypeptide containing a region of Gal4 which interacts
CC   with Gal11P into a cell containing Gal11P and identifying members of the
CC   library that interact with the target from activation of transcription.
CC   Such constructs are used to activate transcription in a cell, e.g. for
CC   controlling gene activity, particularly in gene therapy (e.g. recognizing
CC   a site close to a selected therapeutic gene). Transcription can be
CC   activated without blocking other transcriptional activators. They
CC   probably act by interacting with a component of the RNA polymerase II
CC   holoenzyme, Gal11, the strongest known yeast activator, which provides a
CC   more sensitive assay allowing detection of even weak protein-protein
CC   interactions. Such activators do not create toxicity problems even when
CC   overexpressed
XX
SQ   Sequence 6 AA;

  Query Match             57.1%;  Score 20;  DB 2;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFTL 5
              ||| |
Db          1 TDFLL 5


RESULT 10
AAP82200
ID   AAP82200 standard; protein; 5 AA.
XX
AC   AAP82200;
XX
DT   25-OCT-1990  (first entry)
XX
DE   Example of peptide 2 for treatment of schizophrenia or psoriasis.
XX
KW   schizophrenia; psoriasis; vasoactive intestinal polypeptide.
XX
OS   Synthetic.
XX
PN   SE8700125-A.
XX
PD   16-JUL-1988.
XX
PF   15-JAN-1987;   87SE-00000125.
XX
PR   15-JAN-1987;   87SE-00000125.
XX
PA   (WETT/) WETTERBERG.
XX
PI   Wetterberg L;
XX
DR   WPI; 1988-328337/46.
XX
PT   Short peptide(s) for treatment of psoriasis and schizophrenia - comprise
PT   vaso-active intestinal polypeptide, peptide T or short peptide with five
PT   aminoacid(s).
XX
PS   Claim 1; Page 5; 9pp; Swedish.
XX
CC   Specific example of pentapeptide of the general formula of AAP82197.
CC   These peptides can be administered intravenously, topically or perorally
CC   to relieve the symptoms of psoriasis or schizophrenia. Amino acids at
CC   posns 2 and 3 can be any residue but Asp is preferred at posn 3. See also
CC   AAP82196-9
XX
SQ   Sequence 5 AA;

  Query Match             54.3%;  Score 19;  DB 1;  Length 5;
  Best Local Similarity   75.0%;  Pred. No. 1.4e+06;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TDFT 4
              ||:|
Db          2 TDYT 5


RESULT 11
ADE65091
ID   ADE65091 standard; peptide; 5 AA.
XX
AC   ADE65091;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Corticotropin-releasing factor-2 polypeptide, SEQ ID No 394.
XX
KW   corticotropin-releasing factor-2; CRF2; myopathic; osteopathic;
KW   hypotensive; cardiant; vasotropic; antimigraine; cerebroprotective;
KW   nootropic; neuroprotective; anorectic; antidiabetic; analgesic;
KW   antiallergic; tranquilizer; anxiolytic; antidepressant; antiarthritic;
KW   gene therapy.
XX
OS   Unidentified.
XX
PN   WO2003062277-A1.
XX
PD   31-JUL-2003.
XX
PF   16-JAN-2003; 2003WO-US001454.
```

```
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Desmet J,  Hufton S,  Hoogenboom H,  Sablon E;
XX
DR   WPI; 2000-665002/64.
XX
PT   Scaffold composed of single-chain polypeptide having beta sandwich
PT   architecture carrying new and randomized peptide sequences useful as
PT   supporting framework and carrying antigen- or receptor binding fragments.
XX
PS   Disclosure; Page 15; 68pp; English.
XX
CC   The present invention is concerned with producing scaffold proteins based
CC   upon the human CTLA-4 SCA domain. These scaffold proteins can be used as
CC   a scaffold to bind antigen- or receptor-binding fragments. These can be
CC   used in the treatment of diseases such as cancer, atherosclerosis,
CC   thrombosis, osteoporosis, rheumatoid arthritis and diabetic retinopathy.
CC   Sequences AAB29930-B29939 were used in the production of the proteins of
CC   the invention
XX
SQ   Sequence 7 AA;

  Query Match             74.3%;  Score 26;  DB 3;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DFTLTI 7
              ||||:|
Db          1 DFTLSI 6


RESULT 7
AAY14399
ID   AAY14399 standard; peptide; 7 AA.
XX
AC   AAY14399;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Peptide CDR-H1-7(Y3F) derived from anti-HCV protease MAb 8D4.
XX
KW   Complementarity determining region; CDR; monoclonal antibody; MAb;
KW   hepatitis C virus; HCV; protease; binding site.
XX
OS   Synthetic.
XX
PN   JP11127861-A.
XX
PD   18-MAY-1999.
XX
PF   29-OCT-1997;   97JP-00297451.
XX
PR   29-OCT-1997;   97JP-00297451.
XX
PA   (NIHA ) JAPAN ENERGY CORP.
XX
DR   WPI; 1999-350322/30.
XX
PT   Neutralized antibody partial peptide derived from hepatitis C virus -
PT   useful for inhibiting Hepatitis C Virus (HCV) serine protease activity.
XX
PS   Example 1; Page 24; 32pp; Japanese.
XX
CC   This sequence corresponds to a peptide (CDR-H1-7; AAY14403) derived from
CC   the sequence of the heavy chain variable region complementarity
CC   determining region (CDR)-1 of the anti-hepatitis C virus (HCV) Ser/Thr
CC   protease monoclonal antibody (MAb) 8D4 protein. The peptide has a Tyr to
CC   Phe amino acid substitution at position 3 compared to the CDR-H1-7
CC   peptide. The invention relates to the use of partial peptides (AAY14348-
CC   Y14353) from the MAb 8D4 for inhibiting HCV serine protease activity
XX
SQ   Sequence 7 AA;

  Query Match             60.0%;  Score 21;  DB 2;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFTL 5
              ||| |
Db          1 TDFVL 5


RESULT 8
AAU85454
ID   AAU85454 standard; peptide; 5 AA.
XX
AC   AAU85454;
XX
DT   21-MAY-2002  (first entry)
XX
DE   Human colon specific polypeptide antibody binding site #31.
XX
KW   Human; colon specific gene; CSG; cytostatic; metastasis;
KW   colon cancer staging; antibody binding site.
XX
OS   Homo sapiens.
XX
PN   WO200206515-A2.
XX
PD   24-JAN-2002.
XX
PF   17-JUL-2001; 2001WO-US022454.
XX
PR   17-JUL-2000; 2000US-00618596.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Macina RA,  Sun Y;
XX
DR   WPI; 2002-171815/22.
XX
PT   Diagnosing, staging or monitoring colon cancer involves determining a
PT   colon specific gene in cells, tissues or body fluids in patient, and
PT   comparing it with levels of the gene from a normal human control.
XX
PS   Disclosure; Page 21; 52pp; English.
XX
CC   The invention relates to diagnosing the presence of colon cancer,
CC   metastases of colon cancer, staging colon cancer, monitoring colon cancer
CC   for the onset of metastasis or monitoring a change in stage of colon
CC   cancer in a patient. The method involves determining a colon specific
CC   gene (CSG) in cells, tissues or bodily fluids and comparing it with
CC   levels of CSG in cells, tissues or bodily fluids from a normal human
CC   control. Colon cancer can be treated by administering a molecule which
CC   down regulates the expression or activity of CSG. An immune response
CC   against a target cell expressing CSG can be induced by delivering an
CC   immunologically stimulatory amount of a CSG protein to a patient, so that
CC   an immune response is mounted. Therapeutic agents are useful for imaging
CC   colon cancer in a patient by administering an agent labelled with
CC   paramagnetic ions or a radioisotope to the patient. They are also useful
CC   for preventing the onset of colon cancer, and in diagnosis and treatment
CC   of the disease. Sequences AAU85424-AAU85502 represent human colon
CC   specific protein antibody binding sites used in the method of the
CC   invention
XX
SQ   Sequence 5 AA;

  Query Match             57.1%;  Score 20;  DB 5;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FTLT 6
              ||||
Db          2 FTLT 5
```

```
AAB30075
ID   AAB30075 standard; peptide; 7 AA.
XX
AC   AAB30075;
XX
DT   09-FEB-2001  (first entry)
XX
DE   Scaffold protein SCA S4 peptide SEQ ID NO: 136.
XX
KW   Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW   SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW   diabetic retinopathy; atherosclerosis.
XX
OS   Synthetic.
XX
PN   WO200060070-A1.
XX
PD   12-OCT-2000.
XX
PF   01-APR-1999;   99WO-EP002283.
XX
PR   01-APR-1999;   99WO-EP002283.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Desmet J,  Hufton S,  Hoogenboom H,  Sablon E;
XX
DR   WPI; 2000-665002/64.
XX
PT   Scaffold composed of single-chain polypeptide having beta sandwich
PT   architecture carrying new and randomized peptide sequences useful as
PT   supporting framework and carrying antigen- or receptor binding fragments.
XX
PS   Disclosure; Page 15; 68pp; English.
XX
CC   The present invention is concerned with producing scaffold proteins based
CC   upon the human CTLA-4 SCA domain. These scaffold proteins can be used as
CC   a scaffold to bind antigen- or receptor-binding fragments. These can be
CC   used in the treatment of diseases such as cancer, atherosclerosis,
CC   thrombosis, osteoporosis, rheumatoid arthritis and diabetic retinopathy.
CC   Sequences AAB29930-B29939 were used in the production of the proteins of
CC   the invention
XX
SQ   Sequence 7 AA;

  Query Match             77.1%;  Score 27;  DB 3;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DFTLTI 7
              |:||||
Db          1 DYTLTI 6


RESULT 5
AAY40736
ID   AAY40736 standard; peptide; 7 AA.
XX
AC   AAY40736;
XX
DT   01-DEC-1999  (first entry)
XX
DE   S4 derivative #10, beta strand of scaffold protein structure.
XX
KW   Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW   tumour; chemotherapeutic agent.
XX
OS   Synthetic.
XX
PN   EP947582-A1.
XX
PD   06-OCT-1999.
XX
PF   31-MAR-1998;   98EP-00870065.
XX
PR   31-MAR-1998;   98EP-00870065.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Desmet J,  Hufton S,  Hoogenboom H,  Sablon E;
XX
DR   WPI; 1999-542958/46.
XX
PT   New scaffold protein, useful for stabilizing antigens used as vaccines.
XX
PS   Disclosure; Page 6; 105pp; English.
XX
CC   Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC   S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC   beta strand peptide which forms part of a beta sheet. Peptides (AAY40601-
CC   Y40609) together form a single-chain scaffold protein which contains at
CC   least 1 disulfide bond, contains less than 10% alpha helix and contains
CC   at least 6 beta-strands. The scaffold protein is constructed of beta
CC   strands S1-S6, and may also include beta strands A1-A3, or any
CC   functionally equivalent derivative of these sequences. The beta strands
CC   form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC   the next by hydrogen bonds, which generate a beta sandwich architecture.
CC   If the additional beta strands A1-A3 are included in the structure the
CC   scaffold is constructed of two beta sheets, with the structures
CC   A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC   other via amino acid loops, where at least one of the loops binds to a
CC   receptor or antigen. The scaffold protein is used to stabilize antigens
CC   or whole proteins such as receptors, or their fragments. It may be used
CC   to bind two separate molecules. For example, one surface of the scaffold
CC   may be bound to a protein which binds to a tumour antigen. This will
CC   target the complex to tumour cells. Another surface may be bound to a
CC   cytotoxic molecule or an autoimmune antibody which may then kill the
CC   tumour cells. Therefore the scaffold protein may be used to target
CC   chemotherapeutic agents to specific cells. It may also be used to
CC   stabilize individual peptides in a peptide library and may be used in
CC   diagnostic techniques, and to stabilize antigens used as vaccines
XX
SQ   Sequence 7 AA;

  Query Match             74.3%;  Score 26;  DB 2;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DFTLTI 7
              ||||:|
Db          1 DFTLSI 6


RESULT 6
AAB30074
ID   AAB30074 standard; peptide; 7 AA.
XX
AC   AAB30074;
XX
DT   09-FEB-2001  (first entry)
XX
DE   Scaffold protein SCA S4 peptide SEQ ID NO: 135.
XX
KW   Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW   SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW   diabetic retinopathy; atherosclerosis.
XX
OS   Synthetic.
XX
PN   WO200060070-A1.
XX
PD   12-OCT-2000.
XX
PF   01-APR-1999;   99WO-EP002283.
XX
PR   01-APR-1999;   99WO-EP002283.
```

CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines
XX
SQ Sequence 7 AA;

Query Match 85.7%; Score 30; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DFTLTI 7
||||||
Db 1 DFTLTI 6

RESULT 2
AAB30076
ID AAB30076 standard; peptide; 7 AA.
XX
AC AAB30076;
XX
DT 09-FEB-2001 (first entry)
XX
DE Scaffold protein SCA S4 peptide SEQ ID NO: 137.
XX
KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.
XX
OS Synthetic.
XX
PN WO200060070-A1.
XX
PD 12-OCT-2000.
XX
PF 01-APR-1999; 99WO-EP002283.
XX
PR 01-APR-1999; 99WO-EP002283.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 2000-665002/64.
XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding fragments.
XX
PS Disclosure; Page 15; 68pp; English.
XX
CC The present invention is concerned with producing scaffold proteins based
CC upon the human CTLA-4 SCA domain. These scaffold proteins can be used as
CC a scaffold to bind antigen- or receptor-binding fragments. These can be
CC used in the treatment of diseases such as cancer, atherosclerosis,
CC thrombosis, osteoporosis, rheumatoid arthritis and diabetic retinopathy.
CC Sequences AAB29930-B29939 were used in the production of the proteins of
CC the invention
XX
SQ Sequence 7 AA;

Query Match 85.7%; Score 30; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DFTLTI 7
||||||
Db 1 DFTLTI 6

RESULT 3
AAY40737
ID AAY40737 standard; peptide; 7 AA.
XX
AC AAY40737;
XX
DT 01-DEC-1999 (first entry)
XX
DE S4 derivative #11, beta strand of scaffold protein structure.
XX
KW Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.
XX
OS Synthetic.
XX
PN EP947582-A1.
XX
PD 06-OCT-1999.
XX
PF 31-MAR-1998; 98EP-00870065.
XX
PR 31-MAR-1998; 98EP-00870065.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 1999-542958/46.
XX
PT New scaffold protein, useful for stabilizing antigens used as vaccines.
XX
PS Disclosure; Page 6; 105pp; English.
XX
CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides (AAY40601-
CC Y40609) together form a single-chain scaffold protein which contains at
CC least 1 disulfide bond, contains less than 10% alpha helix and contains
CC at least 6 beta-strands. The scaffold protein is constructed of beta
CC strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines
XX
SQ Sequence 7 AA;

Query Match 77.1%; Score 27; DB 2; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DFTLTI 7
|:||||
Db 1 DYTLTI 6

RESULT 4

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:40:31 ; Search time 45.6667 Seconds
(without alignments)
43.310 Million cell updates/sec

Title: US-09-712-819C-5
Perfect score: 35
Sequence: 1 TDFTLTI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 92273

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	85.7	7	2	AAY40738	Aay40738 S4 deriva
2	30	85.7	7	3	AAB30076	Aab30076 Scaffold
3	27	77.1	7	2	AAY40737	Aay40737 S4 deriva
4	27	77.1	7	3	AAB30075	Aab30075 Scaffold
5	26	74.3	7	2	AAY40736	Aay40736 S4 deriva
6	26	74.3	7	3	AAB30074	Aab30074 Scaffold
7	21	60.0	7	2	AAY14399	Aay14399 Peptide C
8	20	57.1	5	5	AAU85454	Aau85454 Human col
9	20	57.1	6	2	AAW31467	Aaw31467 Transcrip
10	19	54.3	5	1	AAP82200	Aap82200 Example o
11	19	54.3	5	7	ADE65091	Ade65091 Corticotr
12	19	54.3	5	7	ADE51377	Ade51377 CRF2 non-
13	19	54.3	6	2	AAY06466	Aay06466 Epitope t
14	19	54.3	6	2	AAE16558	Aae16558 Epitope t
15	19	54.3	6	3	AAY77708	Aay77708 AU5 pepti
16	19	54.3	6	4	AAB59859	Aab59859 AU5 pepti
17	19	54.3	6	4	AAE13076	Aae13076 Epitope t
18	19	54.3	6	4	AAM51422	Aam51422 Integrin
19	19	54.3	6	4	AAB97355	Aab97355 AU5 epito
20	19	54.3	6	5	AAE24897	Aae24897 AU5 pepti
21	19	54.3	6	5	ABG32853	Abg32853 Epitope t
22	19	54.3	6	7	ADD67264	Add67264 AU5 epito
23	19	54.3	6	7	ADE65099	Ade65099 Corticotr
24	19	54.3	6	7	ADE51385	Ade51385 CRF2 non-
25	19	54.3	7	3	AAY52584	Aay52584 Amaranthu
26	18	51.4	5	6	ABR55418	Abr55418 Amino aci
27	18	51.4	6	2	AAW75358	Aaw75358 Hexapepti
28	18	51.4	6	2	AAW75290	Aaw75290 Hexapepti
29	18	51.4	7	2	AAR09409	Aar09409 LFA-1 alp
30	18	51.4	7	2	AAW58711	Aaw58711 Tryptic 4
31	18	51.4	7	2	AAY14403	Aay14403 Peptide C
32	18	51.4	7	4	AAG64484	Aag64484 Antihepat
33	17	48.6	4	5	ABB55625	Abb55625 Mutated p
34	17	48.6	4	5	ABB55651	Abb55651 Mutated p
35	17	48.6	5	3	AAY51466	Aay51466 AAV VP3 d
36	17	48.6	5	3	AAB17215	Aab17215 IL-1 anta
37	17	48.6	5	3	AAB52195	Aab52195 Human ant
38	17	48.6	5	5	ABB72462	Abb72462 Interleuk
39	17	48.6	5	5	ADE53380	Ade53380 FEN-1 rel
40	17	48.6	5	6	ABU12234	Abu12234 Streptoco
41	17	48.6	5	6	ABU12235	Abu12235 Streptoco
42	17	48.6	6	2	AAY06532	Aay06532 Epidermal
43	17	48.6	6	2	AAY33711	Aay33711 Hepatoma
44	17	48.6	6	3	AAY95393	Aay95393 Human pan
45	17	48.6	6	4	AAB87699	Aab87699 Hepatoma-

ALIGNMENTS

RESULT 1
AAY40738
ID AAY40738 standard; peptide; 7 AA.
XX
AC AAY40738;
XX
DT 01-DEC-1999 (first entry)
XX
DE S4 derivative #12, beta strand of scaffold protein structure.
XX
KW Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.
XX
OS Synthetic.
XX
PN EP947582-A1.
XX
PD 06-OCT-1999.
XX
PF 31-MAR-1998; 98EP-00870065.
XX
PR 31-MAR-1998; 98EP-00870065.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 1999-542958/46.
XX
PT New scaffold protein, useful for stabilizing antigens used as vaccines.
XX
PS Disclosure; Page 6; 105pp; English.
XX
CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides (AAY40601-
CC Y40609) together form a single-chain scaffold protein which contains at
CC least 1 disulfide bond, contains less than 10% alpha helix and contains
CC at least 6 beta-strands. The scaffold protein is constructed of beta
CC strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens

```
RT   "A novel active pentapeptide from chicken brain identified by
RT   antibodies to FMRFamide.";
RL   Nature 305:328-330(1983).
CC   -!- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   GO; GO:0007218; P:neuropeptide signaling pathway; TAS.
KW   Neuropeptide; Amidation.
FT   MOD_RES       5       5       AMIDATION.
SQ   SEQUENCE   5 AA;  645 MW;  69D4073767400000 CRC64;

  Query Match             17.1%;  Score 6;  DB 13;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 F 3
              |
Db          5 F 5


RESULT 15
P83569
ID   P83569      PRELIMINARY;      PRT;     6 AA.
AC   P83569;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Sperm attracting peptide SepSAP.
OS   Sepia officinalis (Common cuttlefish).
OC   Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC   Decapodiformes; Sepioidea; Sepiidae; Sepia.
OX   NCBI_TaxID=6610;
RN   [1]
RP   SEQUENCE, FUNCTION, DEVELOPMENTAL STAGE, MASS SPECTROMETRY, AND
RP   AMIDATION.
RC   TISSUE=Egg;
RX   PubMed=12207899;
RA   Zatylny C., Marvin L., Gagnon J., Henry J.;
RT   "Fertilization in Sepia officinalis: the first mollusk sperm-
RT   attracting peptide.";
RL   Biochem. Biophys. Res. Commun. 296:1186-1193(2002).
CC   -!- FUNCTION: ATTRACTS SPERM INCREASING THE CHANCES OF GAMETE
CC       COLLISION.
CC   -!- DEVELOPMENTAL STAGE: FIRST APPEARS IN THE OVARIAN FOLLICLES FURING
CC       VITELLOGENESIS. ACCUMULATES IN THE OOCYTES BEFORE BEING SECRETED
CC       DURING FERTILIZATION. EXPRESSION CONTINUES IN THE EMBEDDED OOCYTE.
CC       ACCUMULATES IN THE EGG CAPSULE AFTER FERTILIZATION.
CC   -!- MASS SPECTROMETRY: MW=596.6; METHOD=MALDI.
KW   Amidation.
FT   MOD_RES       6       6       AMIDATION.
SQ   SEQUENCE   6 AA;  597 MW;  72C8676AA0470000 CRC64;

  Query Match             17.1%;  Score 6;  DB 5;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 D 2
              |
Db          3 D 3


Search completed: June  1, 2004, 11:18:36
Job time : 31.3333 secs
```

```
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   88 kDa protein (Fragment).
OS   Bacillus cereus.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1396;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=NCIMB 11796;
RA   Browne N., Dowds B.C.A.;
RL   Submitted (JUL-2001) to Swiss-Prot.
FT   NON_TER       5       5
SQ   SEQUENCE   5 AA;  623 MW;  6B01AAA336F00000 CRC64;

  Query Match             20.0%;  Score 7;  DB 2;  Length 5;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TD 2
              |:
Db          4 TE 5


RESULT 11
Q47029
ID   Q47029      PRELIMINARY;      PRT;     7 AA.
AC   Q47029;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Aad A1 protein (Fragment).
GN   AAD A1.
OS   Enterobacter cloacae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Enterobacter.
OX   NCBI_TaxID=550;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94079349; PubMed=8257126;
RA   Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;
RT   "Analysis of the aac(3)-VIa gene encoding a novel 3-N-
RT   acetyltransferase.";
RL   Antimicrob. Agents Chemother. 37:2074-2079(1993).
DR   EMBL; M88012; AAA16193.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   7 AA;  744 MW;  633862D2C321A030 CRC64;

  Query Match             20.0%;  Score 7;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LT 6
              :|
Db          1 IT 2


RESULT 12
P72081
ID   P72081      PRELIMINARY;      PRT;     7 AA.
AC   P72081;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   3'-methylcephem hydroxylase (Fragment).
GN   CEFF.
OS   Nocardia lactamdurans.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Pseudonocardineae; Pseudonocardiaceae; Amycolatopsis.
OX   NCBI_TaxID=1913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96009872; PubMed=7557411;
RA   Coque J., Perez-Llarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,
RA   Liras P.;
RT   "Characterization of the cmcH genes of Nocardia lactamdurans and
RT   Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
RT   O-carbamoyltransferase for cephamycin biosynthesis.";
RL   Gene 162:21-27(1995).
DR   EMBL; Z21682; CAA79797.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   7 AA;  746 MW;  75A1B2CDD1E681C0 CRC64;

  Query Match             20.0%;  Score 7;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TL 5
              |:
Db          1 TM 2


RESULT 13
O98866
ID   O98866      PRELIMINARY;      PRT;     7 AA.
AC   O98866;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cytochrome b/f subunit IV (Fragment).
OS   Spinacia oleracea (Spinach).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86120353; PubMed=3003688;
RA   Sijben-Mueller G., Hallick R.B., Alt J., Westhoff P., Herrmann R.G.;
RT   "Spinach plastid genes coding for initiation factor IF-1, ribosomal
RT   protein S11 and RNA polymerase alpha-subunit.";
RL   Nucleic Acids Res. 14:1029-1044(1986).
DR   EMBL; X03496; CAA27215.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1       1
SQ   SEQUENCE   7 AA;  907 MW;  644729D77409C420 CRC64;

  Query Match             20.0%;  Score 7;  DB 8;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DF 3
              :|
Db          1 NF 2


RESULT 14
P83308
ID   P83308      PRELIMINARY;      PRT;     5 AA.
AC   P83308;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   FMRFamide-like neuropeptide (LPLRF-amide).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Brain;
RX   PubMed=6137771;
RA   Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;
```

```
RESULT 6
P82445
ID   P82445      PRELIMINARY;      PRT;     7 AA.
AC   P82445;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   10 kDa cell wall protein (Fragment).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. PETIT HAVANA;
RA   Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
RA   Wojtaszek P., Bolwell G.P.;
RT   "Proteomic study of secondary cell wall proteins from transformed
RT   tobacco culture.";
RL   Planta 0:0-0(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL.
CC   -!- TISSUE SPECIFICITY: XYLEM.
DR   GO; GO:0005618; C:cell wall; IEA.
KW   Cell wall.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  758 MW;  69D2C1E862D1B2A0 CRC64;

  Query Match             25.7%;  Score 9;  DB 10;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 1e+06;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LTI 7
              :|:
Db          1 VTV 3


RESULT 7
Q8JE81
ID   Q8JE81      PRELIMINARY;      PRT;     7 AA.
AC   Q8JE81;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Truncated pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=4874;
RX   MEDLINE=22056123; PubMed=12060770;
RA   Beerenwinkel N., Schmidt B., Walter H., Kaiser R., Lengauer T.,
RA   Hoffmann D., Korn K., Selbig J.;
RT   "Diversity and complexity of HIV-1 drug resistance: A bioinformatics
RT   approach to predicting phenotype from genotype.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:8271-8276(2002).
DR   EMBL; AF347267; AAK32344.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  885 MW;  76C37731A046C700 CRC64;

  Query Match             25.7%;  Score 9;  DB 15;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TL 5
              ||
Db          4 TL 5


RESULT 8
P83533
ID   P83533      PRELIMINARY;      PRT;     6 AA.
AC   P83533;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Unknown protein from 2D-page (Fragment).
OS   Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1625;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=DSM 20451;
RX   PubMed=12112860;
RA   Drews O., Weiss W., Reil G., Parlar H., Wait R., Goerg A.;
RT   "High pressure effects step-wise altered protein expression in
RT   Lactobacillus sanfranciscensis.";
RL   Proteomics 2:765-774(2002).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS UNKNOWN
CC       PROTEIN IS: 15 KDA.
FT   NON_TER       1      1
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  590 MW;  6DDDD452D1AAC000 CRC64;

  Query Match             22.9%;  Score 8;  DB 2;  Length 6;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 TI 7
              |:
Db          2 TV 3


RESULT 9
P70804
ID   P70804      PRELIMINARY;      PRT;     7 AA.
AC   P70804;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AlgT protein (Fragment).
GN   ALGT.
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E;
RX   MEDLINE=96427318; PubMed=8830682;
RA   Rehm B.H.A., Ertesvag H., Valla S.;
RT   "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is
RT   part of an alg gene cluster physically organized in a manner similar
RT   to that in Pseudomonas aeruginosa.";
RL   J. Bacteriol. 178:5884-5889(1996).
DR   EMBL; X87973; CAA61230.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  684 MW;  71B5A5A5A2D1AED0 CRC64;

  Query Match             22.9%;  Score 8;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 TI 7
              |:
Db          2 TV 3


RESULT 10
P83073
ID   P83073      PRELIMINARY;      PRT;     5 AA.
AC   P83073;
```

```
SQ   SEQUENCE   7 AA;  874 MW;  72D1A9DB5041A6F0 CRC64;

  Query Match             34.3%;  Score 12;  DB 5;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FTL 5
              ||:
Db          5 FTV 7


RESULT 2
Q54248
ID   Q54248      PRELIMINARY;      PRT;     7 AA.
AC   Q54248;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RplO protein (Fragment).
GN   RPLO.
OS   Streptomyces griseus.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1911;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2-3-11;
RX   MEDLINE=20011291; PubMed=10542330;
RA   Poehling S., Piepersberg W., Wehmeier U.F.;
RT   "Analysis and regulation of the sec Y gene from Streptomyces griseus
RT   N2-3-11 and interaction of the SecY protein with the SecA protein.";
RL   Biochim. Biophys. Acta 1447:298-302(1999).
DR   EMBL; X95915; CAA65160.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  760 MW;  72C72B01B2D1B2A0 CRC64;

  Query Match             31.4%;  Score 11;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TLT 6
              |:|
Db          2 TVT 4


RESULT 3
P83530
ID   P83530      PRELIMINARY;      PRT;     7 AA.
AC   P83530;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Unknown protein from 2D-page (Fragment).
OS   Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1625;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=DSM 20451;
RX   PubMed=12112860;
RA   Drews O., Weiss W., Reil G., Parlar H., Wait R., Goerg A.;
RT   "High pressure effects step-wise altered protein expression in
RT   Lactobacillus sanfranciscensis.";
RL   Proteomics 2:765-774(2002).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS UNKNOWN
CC       PROTEIN IS: 15 KDA.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  790 MW;  6AAB02CAA731B2A0 CRC64;

  Query Match             31.4%;  Score 11;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 TLTI 7
              || :
Db          2 TLDV 5


RESULT 4
O34028
ID   O34028      PRELIMINARY;      PRT;     7 AA.
AC   O34028;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Catechol-2,3-dioxygenase (Fragment).
GN   PHNE.
OS   Sphingomonas chungbukensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC   Sphingomonadaceae; Sphingomonas.
OX   NCBI_TaxID=56193;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DJ77;
RA   Kim Y.-C.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U88298; AAB66311.1; -.
DR   GO; GO:0016702; F:oxidoreductase activity, acting on single d. . .; IEA.
KW   Dioxygenase.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  868 MW;  71A452D1A699D460 CRC64;

  Query Match             28.6%;  Score 10;  DB 2;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 1e+06;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LTI 7
              :|:
Db          3 MTV 5


RESULT 5
O07354
ID   O07354      PRELIMINARY;      PRT;     7 AA.
AC   O07354;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NifK (Fragment).
GN   NIFK.
OS   Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothece PCC 8801).
OC   Bacteria; Cyanobacteria; Chroococcales; Cyanothece.
OX   NCBI_TaxID=41431;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RF-1;
RX   MEDLINE=99231861; PubMed=10217509;
RA   Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RT   "Organization and expression of nitrogen-fixation genes in the aerobic
RT   nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
RT   RF-1.";
RL   Microbiology 145:743-753(1999).
DR   EMBL; AF003700; AAC35193.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  849 MW;  7412C72AA9D5B030 CRC64;

  Query Match             25.7%;  Score 9;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 FTL 5
              | |
Db          3 FDL 5
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:53:46 ; Search time 31.3333 Seconds
(without alignments)
70.488 Million cell updates/sec

Title: US-09-712-819C-5
Perfect score: 35
Sequence: 1 TDFTLTI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 74

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	12	34.3	7	5	Q7Z1C0	Q7z1c0 caenorhabdi
2	11	31.4	7	2	Q54248	Q54248 streptomyce
3	11	31.4	7	2	P83530	P83530 lactobacill
4	10	28.6	7	2	O34028	O34028 sphingomona
5	9	25.7	7	2	O07354	O07354 synechococc
6	9	25.7	7	10	P82445	P82445 nicotiana t
7	9	25.7	7	15	Q8JE81	Q8je81 human immun
8	8	22.9	6	2	P83533	P83533 lactobacill
9	8	22.9	7	2	P70804	P70804 azotobacter
10	7	20.0	5	2	P83073	P83073 bacillus ce
11	7	20.0	7	2	Q47029	Q47029 enterobacte
12	7	20.0	7	2	P72081	P72081 nocardia la
13	7	20.0	7	8	O98866	O98866 spinacia ol
14	6	17.1	5	13	P83308	P83308 gallus gall
15	6	17.1	6	5	P83569	P83569 sepia offic
16	6	17.1	6	10	P82541	P82541 spinacia ol
17	6	17.1	7	2	O50556	O50556 actinobacil
18	6	17.1	7	2	Q8KMS9	Q8kms9 enterobacte
19	6	17.1	7	4	Q15903	Q15903 homo sapien
20	6	17.1	7	6	Q28742	Q28742 oryctolagus
21	6	17.1	7	8	Q8MFY6	Q8mfy6 taraxacum (
22	6	17.1	7	10	O49223	O49223 glycine max
23	6	17.1	7	10	Q9C5B3	Q9c5b3 arabidopsis
24	6	17.1	7	11	Q8K3H6	Q8k3h6 rattus norv
25	6	17.1	7	11	Q63480	Q63480 rattus norv
26	6	17.1	7	11	O55184	O55184 rattus norv
27	6	17.1	7	12	Q9YQ10	Q9yq10 transmissib
28	6	17.1	7	13	Q8JJ20	Q8jj20 gallus gall
29	5	14.3	6	10	P82181	P82181 spinacia ol
30	5	14.3	6	10	P82182	P82182 spinacia ol
31	5	14.3	7	2	Q8KMS3	Q8kms3 klebsiella
32	5	14.3	7	2	Q47505	Q47505 escherichia
33	5	14.3	7	3	P83492	P83492 bionectria
34	5	14.3	7	8	Q95945	Q95945 saccharomyc
35	5	14.3	7	10	P93233	P93233 lycopersico
36	5	14.3	7	12	Q67113	Q67113 influenzavi
37	5	14.3	7	12	Q65578	Q65578 bovine herp
38	5	14.3	7	13	O42564	O42564 fugu rubrip
39	5	14.3	7	15	Q07624	Q07624 rous sarcom
40	4	11.4	4	5	P83568	P83568 sepia offic
41	4	11.4	4	11	Q08433	Q08433 rattus sp.
42	4	11.4	7	2	Q8GL12	Q8gl12 borrelia bu
43	4	11.4	7	2	Q8GL04	Q8gl04 borrelia bu
44	4	11.4	7	2	Q8GL00	Q8gl00 borrelia bu
45	4	11.4	7	4	Q8NHH7	Q8nhh7 homo sapien

ALIGNMENTS

RESULT 1
Q7Z1C0
ID Q7Z1C0 PRELIMINARY; PRT; 7 AA.
AC Q7Z1C0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein W01B11.6.
GN W01B11.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Bradshaw H., Graves T., Blair T.;
RT "The sequence of C. elegans cosmid W01B11.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF043704; AAO38592.1; -.
KW Hypothetical protein.

```
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Multigene family.
SQ   SEQUENCE   7 AA;  796 MW;  672879CDCB476B70 CRC64;

  Query Match             22.9%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 FTL 5
              | |
Db          5 FGL 7


RESULT 14
ALL4_CARMA
ID   ALL4_CARMA     STANDARD;      PRT;     7 AA.
AC   P81807;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 4.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Multigene family.
SQ   SEQUENCE   7 AA;  782 MW;  672879CDCB476AC0 CRC64;

  Query Match             22.9%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 FTL 5
              | |
Db          5 FGL 7


RESULT 15
ALL5_CARMA
ID   ALL5_CARMA     STANDARD;      PRT;     7 AA.
AC   P81808;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 5.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       7       7       AMIDATION.
SQ   SEQUENCE   7 AA;  781 MW;  672879CDCB476420 CRC64;

  Query Match             22.9%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 FTL 5
              | |
Db          5 FGL 7


Search completed: June  1, 2004, 11:16:49
Job time : 7.66667 secs
```

```
Qy          3 FTL 5
              | |
Db          3 FGL 5


RESULT 10
PSK_DAUCA
ID   PSK_DAUCA      STANDARD;      PRT;     5 AA.
AC   P58261;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Phytosulfokine-alpha (PSK-alpha) [Contains: Phytosulfokine-beta (PSK-
DE   beta)].
OS   Daucus carota (Carrot).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   campanulids; Apiales; Apiaceae; Apioideae; Scandiceae; Daucinae;
OC   Daucus.
OX   NCBI_TaxID=4039;
RN   [1]
RP   SEQUENCE, AND IDENTIFICATION BY MASS SPECTROMETRY.
RC   STRAIN=cv. US-Harumakigosun;
RX   MEDLINE=20212743; PubMed=10750705;
RA   Hanai H., Matsuno T., Yamamoto M., Matsubayashi Y., Kobayashi T.,
RA   Kamada H., Sakagami Y.;
RT   "A secreted peptide growth factor, phytosulfokine, acting as a
RT   stimulatory factor of carrot somatic embryo formation.";
RL   Plant Cell Physiol. 41:27-32(2000).
CC   -!- FUNCTION: In presence of 2,4-D, stimulates proliferation of the
CC       cells, but does not stimulate differentiation into the somatic
CC       embryos.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: Sulfation is important for activity and for the binding to a
CC       putative membrane receptor (By similarity).
CC   -!- SIMILARITY: Belongs to the phytosulfokine family.
KW   Growth factor; Sulfation.
FT   PEPTIDE        1       4       PHYTOSULFOKINE-BETA.
FT   MOD_RES        1       1       SULFATION.
FT   MOD_RES        3       3       SULFATION.
SQ   SEQUENCE   5 AA;  687 MW;  76C1BB504B300000 CRC64;

  Query Match             22.9%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FT 4
              :|
Db          3 YT 4


RESULT 11
RE21_LITRU
ID   RE21_LITRU     STANDARD;      PRT;     5 AA.
AC   P82071;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Rubellidin 2.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Skin secretion;
RA   Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australin red tree frog
RT   'Litoria rubella'. The skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC   -!- MASS SPECTROMETRY: MW=626; METHOD=FAB.
KW   Amphibian defense peptide.
SQ   SEQUENCE   5 AA;  626 MW;  6DD9C9CB10300000 CRC64;

  Query Match             22.9%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DF 3
              :|
Db          2 EF 3


RESULT 12
ALL2_CARMA
ID   ALL2_CARMA     STANDARD;      PRT;     7 AA.
AC   P81805;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 2.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES        7       7       AMIDATION (POTENTIAL).
SQ   SEQUENCE   7 AA;  770 MW;  672879CDCB5DDB70 CRC64;

  Query Match             22.9%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 FTL 5
              | |
Db          5 FGL 7


RESULT 13
ALL3_CARMA
ID   ALL3_CARMA     STANDARD;      PRT;     7 AA.
AC   P81806;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 3.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
```

```
Qy          3 FT 4
              ||
Db          4 FT 5


RESULT 6
EI04_LITRU
ID   EI04_LITRU     STANDARD;      PRT;     5 AA.
AC   P82100;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Electrin 4.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Skin secretion;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litori electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:639-645(1999).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Skin.
KW   Amphibian defense peptide; Amidation.
FT   MOD_RES       5       5       AMIDATION.
SQ   SEQUENCE   5 AA;  616 MW;  61F2D1A059A00000 CRC64;

  Query Match             28.6%;  Score 10;  DB 1;  Length 5;
  Best Local Similarity   33.3%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LTI 7
              :|:
Db          2 ITV 4


RESULT 7
CCF1_ENTFA
ID   CCF1_ENTFA     STANDARD;      PRT;     7 AA.
AC   P20104;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cCF10.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=89008313; PubMed=3139658;
RA   Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,
RA   Adsit J.C., Dunny G.M., Suzuki A.;
RT   "Structure of cCF10, a peptide sex pheromone which induces
RT   conjugative transfer of the Streptococcus faecalis tetracycline
RT   resistance plasmid, pCF10.";
RL   J. Biol. Chem. 263:14574-14578(1988).
CC   -!- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       HEMOLYSIN PLASMID PCF10.
DR   PIR; A30812; A30812.
KW   Pheromone.
SQ   SEQUENCE   7 AA;  790 MW;  72C9D2C731B2C740 CRC64;

  Query Match             25.7%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TL 5
              ||
Db          3 TL 4


RESULT 8
CIA_ENTFA
ID   CIA_ENTFA      STANDARD;      PRT;     7 AA.
AC   P11932;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cAM373 (Clumping-inducing agent) (CIA).
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=87005252; PubMed=3093276;
RA   Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
RA   White B.A., An F.Y., Clewell D.B., Suzuki A.;
RT   "Isolation and structure of the Streptococcus faecalis sex pheromone,
RT   cAM373.";
RL   FEBS Lett. 206:69-72(1986).
CC   -!- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
CC       HARBORING PAM373.
CC   -!- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
CC       SPECIFICITY OF PHEROMONES TO PLASMIDS.
CC   -!- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
DR   PIR; A25269; A25269.
KW   Pheromone.
SQ   SEQUENCE   7 AA;  734 MW;  75BDD72059C05DB0 CRC64;

  Query Match             25.7%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 FTL 5
              | |
Db          3 FIL 5


RESULT 9
AL14_CARMA
ID   AL14_CARMA     STANDARD;      PRT;     5 AA.
AC   P81817;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 14.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       5       5       AMIDATION (POTENTIAL).
SQ   SEQUENCE   5 AA;  586 MW;  672879D5AB300000 CRC64;

  Query Match             22.9%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC   -!- MASS SPECTROMETRY: MW=598; METHOD=FAB.
KW   Amphibian defense peptide.
SQ   SEQUENCE   5 AA;  598 MW;  6DD9C9CAB2A00000 CRC64;

  Query Match             34.3%;  Score 12;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DF 3
               ||
Db          2 DF 3


RESULT 3
BIOA_CITFR
ID   BIOA_CITFR     STANDARD;      PRT;     5 AA.
AC   P13071;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
DE   (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA
DE   aminotransferase) (Fragment).
GN   BIOA.
OS   Citrobacter freundii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Citrobacter.
OX   NCBI_TaxID=546;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89006280; PubMed=2971595;
RA   Shiuan D., Campbell A.;
RT   "Transcriptional regulation and gene arrangement of Escherichia coli,
RT   Citrobacter freundii and Salmonella typhimurium biotin operons.";
RL   Gene 67:203-211(1988).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-
CC       oxononanoate = S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-
CC       diaminononanoate.
CC   -!- COFACTOR: Pyridoxal phosphate.
CC   -!- PATHWAY: Biotin biosynthesis.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SIMILARITY: Belongs to class-III of pyridoxal-phosphate-dependent
CC       aminotransferases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21922; -; NOT_ANNOTATED_CDS.
DR   PIR; I40697; I40697.
DR   InterPro; IPR005814; Aminotrans_3.
DR   PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.
KW   Biotin biosynthesis; Transferase; Aminotransferase;
KW   Pyridoxal phosphate.
FT   NON_TER       5       5
SQ   SEQUENCE   5 AA;  582 MW;  6AAAB1B1A6F00000 CRC64;

  Query Match             31.4%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TD 2
               ||
Db          3 TD 4


RESULT 4
RE31_LITRU
ID   RE31_LITRU     STANDARD;      PRT;     5 AA.
AC   P82072;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Rubellidin 3.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Skin secretion;
RA   Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australin red tree frog
RT   'Litoria rubella'. The skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC   -!- MASS SPECTROMETRY: MW=655; METHOD=FAB.
KW   Amphibian defense peptide; Amidation.
FT   MOD_RES       5       5       AMIDATION.
SQ   SEQUENCE   5 AA;  656 MW;  71A9C9CB10300000 CRC64;

  Query Match             31.4%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FT 4
               ||
Db          4 FT 5


RESULT 5
RE32_LITRU
ID   RE32_LITRU     STANDARD;      PRT;     5 AA.
AC   P82073;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Rubellidin 3.2.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Skin secretion;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litori electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:639-645(1999).
CC   -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
KW   Amphibian defense peptide.
SQ   SEQUENCE   5 AA;  570 MW;  71A9C9C862A00000 CRC64;

  Query Match             31.4%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:41:50 ; Search time 7.66667 Seconds
(without alignments)
47.542 Million cell updates/sec

Title: US-09-712-819C-5
Perfect score: 35
Sequence: 1 TDFTLTI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 88

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	14	40.0	7	1	ALL7_CYDPO	P82158 cydia pomon
2	12	34.3	5	1	RE11_LITRU	P82070 litoria rub
3	11	31.4	5	1	BIOA_CITFR	P13071 citrobacter
4	11	31.4	5	1	RE31_LITRU	P82072 litoria rub
5	11	31.4	5	1	RE32_LITRU	P82073 litoria rub
6	10	28.6	5	1	EI04_LITRU	P82100 litoria rub
7	9	25.7	7	1	CCF1_ENTFA	P20104 enterococcu
8	9	25.7	7	1	CIA_ENTFA	P11932 enterococcu
9	8	22.9	5	1	AL14_CARMA	P81817 carcinus ma
10	8	22.9	5	1	PSK_DAUCA	P58261 daucus caro
11	8	22.9	5	1	RE21_LITRU	P82071 litoria rub
12	8	22.9	7	1	ALL2_CARMA	P81805 carcinus ma
13	8	22.9	7	1	ALL3_CARMA	P81806 carcinus ma
14	8	22.9	7	1	ALL4_CARMA	P81807 carcinus ma
15	8	22.9	7	1	ALL5_CARMA	P81808 carcinus ma
16	8	22.9	7	1	FAR1_ASCSU	P31889 ascaris suu
17	7	20.0	4	1	RM01_YEAST	P36515 saccharomyc
18	7	20.0	6	1	FARP_MONEX	P41966 moniezia ex
19	7	20.0	6	1	LOK1_LOCMI	P41491 locusta mig
20	7	20.0	6	1	UN06_CLOPA	P81351 clostridium
21	7	20.0	7	1	FAR1_MACRS	P83274 macrobrachi
22	7	20.0	7	1	FAR1_PROCL	P38499 procambarus
23	7	20.0	7	1	FAR2_PROCL	P38498 procambarus
24	7	20.0	7	1	FAR4_PANRE	P41875 panagrellus
25	7	20.0	7	1	FARB_CALVO	P41866 calliphora
26	7	20.0	7	1	GFRP_MOUSE	P99025 mus musculu
27	7	20.0	7	1	IGAO_DACDE	P06294 dactylium d
28	6	17.1	3	1	LUXE_VIBFI	P24272 vibrio fisc
29	6	17.1	4	1	ACH1_ACHFU	P35904 achatina fu
30	6	17.1	4	1	FAR3_HIRME	P42562 hirudo medi
31	6	17.1	4	1	FAR4_HIRME	P42563 hirudo medi
32	6	17.1	4	1	FFKA_ANTEL	P58705 anthopleura
33	6	17.1	4	1	FLRF_HIRME	P42561 hirudo medi
34	6	17.1	4	1	FLRN_ANTEL	P58707 anthopleura
35	6	17.1	4	1	FMRF_MACNI	P01162 macrocallis
36	6	17.1	4	1	FYRI_ANTEL	P58706 anthopleura
37	6	17.1	4	1	OCP1_OCTMI	P58648 octopus min
38	6	17.1	4	1	OCP3_OCTMI	P58649 octopus min
39	6	17.1	5	1	EI03_LITRU	P82099 litoria rub
40	6	17.1	5	1	FARP_ARTTR	P41853 artioposthi
41	6	17.1	5	1	PAP2_PARMA	P81864 pardachirus
42	6	17.1	5	1	SUGA_ACHDO	P19991 acheta dome
43	6	17.1	5	1	TPIS_CANFA	P54714 canis famil
44	6	17.1	5	1	TRM3_ECOLI	P13973 escherichia
45	6	17.1	5	1	UC22_MAIZE	P80628 zea mays (m

ALIGNMENTS

RESULT 1
ALL7_CYDPO
ID ALL7_CYDPO STANDARD; PRT; 7 AA.
AC P82158;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 7.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricoidea; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -!- SIMILARITY: Belongs to the allatostatin family.
KW Neuropeptide; Amidation.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 873 MW; 672879CABB569350 CRC64;

Query Match 40.0%; Score 14; DB 1; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DFTL 5
|| |
Db 4 DFGL 7

RESULT 2
RE11_LITRU
ID RE11_LITRU STANDARD; PRT; 5 AA.
AC P82070;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rubellidin 1.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC Pelodryadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Skin secretion;
RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australin red tree frog
RT 'Litoria rubella'. The skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";

A;Reference number: S69237; MUID:95139068; PMID:7837271
A;Accession: S69237
A;Molecule type: protein
A;Residues: 1-5 <PET>
A;Experimental source: strain F1, DSM 3639
C;Keywords: cell wall; glycoprotein; heat-stable protein

Query Match 25.7%; Score 9; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          4 TL 5
              ||
Db          2 TL 3
```

RESULT 15
A60986
N-formyl oligopeptide - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 31-Dec-1993
C;Accession: A60986
R;Broom, M.F.; Mellor, D.M.; Chadwick, V.S.
Experientia 45, 1097-1099, 1989
A;Title: Purification and amino acid sequencing of naturally occurring N-formyl-methiony
A;Reference number: A60986; MUID:90092408; PMID:2689204
A;Accession: A60986
A;Molecule type: protein
A;Residues: 1-6 <BRO>
C;Comment: This hexapeptide was the longest of several N-formyl oligopeptides reported.
F;1/Modified site: N-formylmethionine #status experimental

Query Match 25.7%; Score 9; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          3 FTL 5
              | |
Db          3 FIL 5
```

Search completed: June 1, 2004, 11:19:23
Job time : 11.6667 secs

```
  Query Match           28.6%;  Score 10;  DB 2;  Length 6;
  Best Local Similarity  50.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TDFT 4
              :| |
Db          3 SDAT 6


RESULT 9
S09066
globulin IV alpha subunit delta-1 chain, seed - cucurbit (fragments)
N;Alternate names: 11S globulin alpha subunit delta-1 chain
C;Species: Cucurbita sp. (cucurbit)
C;Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 08-Nov-1996
C;Accession: S09066
R;Ohmiya, M.; Hara, I.; Matsubara, H.
Plant Cell Physiol. 21, 157-167, 1980
A;Title: Pumpkin (Cucurbita sp.) seed globulin IV. Terminal sequences of the acidic and
A;Reference number: S09066
A;Accession: S09066
A;Molecule type: protein
A;Residues: 1-6;7 <OHM>

  Query Match           28.6%;  Score 10;  DB 2;  Length 7;
  Best Local Similarity  50.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DFTL 5
              | |:
Db          3 DETI 6


RESULT 10
T13892
cytochrome-c oxidase (EC 1.9.3.1) chain I [imported] - river lamprey mitochondrion (frag
C;Species: mitochondrion Lampetra fluviatilis (river lamprey)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: T13892
R;Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A;Title: The main features of the craniate mitochondrial DNA between the ND1 and the COI
A;Reference number: Z17775; MUID:97398704; PMID:9254918
A;Accession: T13892
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3 <DEL>
A;Cross-references: EMBL:Y09528; NID:g2340016; PIDN:CAA70721.1; PID:g4379123
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

  Query Match           25.7%;  Score 9;  DB 3;  Length 3;
  Best Local Similarity  100.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TL 5
              ||
Db          2 TL 3


RESULT 11
E42364
flagellar protein fliR - Salmonella typhimurium (fragment)
C;Species: Salmonella typhimurium
C;Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 30-Sep-1993
C;Accession: E42364
R;Vogler, A.P.; Homma, M.; Irikura, V.M.; Macnab, R.M.
J. Bacteriol. 173, 3564-3572, 1991
A;Title: Salmonella typhimurium mutants defective in flagellar filament regrowth and seq
A;Reference number: A42364; MUID:91258342; PMID:1646201
A;Accession: E42364
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-5 <VOG>
A;Cross-references: GB:M62408

  Query Match           25.7%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity  100.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TL 5
              ||
Db          3 TL 4


RESULT 12
E60274
major protein antigen MPT63 - Mycobacterium tuberculosis (fragment)
C;Species: Mycobacterium tuberculosis
C;Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 30-Sep-1993
C;Accession: E60274
R;Nagai, S.; Wiker, H.G.; Harboe, M.; Kinomoto, M.
Infect. Immun. 59, 372-382, 1991
A;Title: Isolation and partial characterization of major protein antigens in the cu
A;Reference number: A60274; MUID:91099989; PMID:1898899
A;Accession: E60274
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-5 <NAG>

  Query Match           25.7%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity  25.0%;  Pred. No. 2.8e+05;
  Matches    1;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 FTLT 6
              : :|
Db          2 YPIT 5


RESULT 13
S68326
blood cell protein B - Ascidia ceratodes (fragment)
N;Alternate names: Abcp-B
C;Species: Ascidia ceratodes
C;Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 11-Jun-1999
C;Accession: S68326
R;Taylor, S.W.; Ross, M.M.; Waite, J.H.
Arch. Biochem. Biophys. 324, 228-240, 1995
A;Title: Novel 3,4-di- and 3,4,5-trihydroxyphenylalanine-containing polypeptides fr
A;Reference number: S68325; MUID:96132650; PMID:8554314
A;Accession: S68326
A;Molecule type: protein
A;Residues: 1-5 <TAY>
F;2/Modified site: 3',4',5'-trihydroxyphenylalanine (Tyr) #status experimental
F;4/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental

  Query Match           25.7%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity  50.0%;  Pred. No. 2.8e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DF 3
              |:
Db          1 DY 2


RESULT 14
S69237
surface protein tetrabrachion heavy chain - Staphylothermus marinus (fragment)
C;Species: Staphylothermus marinus
C;Date: 04-Dec-1997 #sequence_revision 04-Dec-1997 #text_change 17-Mar-1999
C;Accession: S69237
R;Peters, J.; Nitsch, M.; Kuehlmorgen, B.; Golbik, R.; Lupas, A.; Kellermann, J.; E
J. Mol. Biol. 245, 385-401, 1995
A;Title: Tetrabrachion: a filamentous archaebacterial surface protein assembly of u
```

RESULT 3
PT0644
T-cell receptor beta chain V-D-J region (111-1G) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0644
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0644
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 34.3%; Score 12; DB 2; Length 5;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TDFT 4
: ||
Db 2 SSFT 5

RESULT 4
B39127
phosphotransferase system enzyme II (EC 2.7.1.69) - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 08-Oct-1999
C;Accession: B39127
R;Hardesty, C.; Ferran, C.; DiRienzo, J.M.
J. Bacteriol. 173, 449-456, 1991
A;Title: Plasmid-mediated sucrose metabolism in Escherichia coli: characterization of sc
rin.
A;Reference number: A39127; MUID:91100329; PMID:1846143
A;Accession: B39127
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-7 <HAR>
A;Cross-references: GB:M38416; NID:g155142; PIDN:AAA98418.1; PID:g155144
C;Keywords: phosphotransferase

Query Match 34.3%; Score 12; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DF 3
||
Db 2 DF 3

RESULT 5
PT0665
T-cell receptor beta chain V-D-J region (121-3BM) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0665
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0665
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 34.3%; Score 12; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DF 3
||
Db 6 DF 7

RESULT 6
I40697
biotin A - Citrobacter freundii (fragment)
C;Species: Citrobacter freundii
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 12-Aug-1996
C;Accession: I40697
R;Shiuan, D.; Campbell, A.
Gene 67, 203-211, 1988
A;Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citr
A;Reference number: I40697; MUID:89006280; PMID:2971595
A;Accession: I40697
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4 <RES>
A;Cross-references: GB:M21922; NID:g144434

Query Match 31.4%; Score 11; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TD 2
||
Db 2 TD 3

RESULT 7
E30608
Ig kappa chain V-III region (Gag) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 16-Aug-1996
C;Accession: E30608
R;Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.
J. Immunol. 142, 3158-3163, 1989
A;Title: Structural and idiotypic characterization of the L chains of human IgM au
A;Reference number: A30601; MUID:89215279; PMID:2496160
A;Accession: E30608
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <GON>
C;Keywords: heterotetramer; immunoglobulin

Query Match 31.4%; Score 11; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DFTLT 6
: ||
Db 1 EIVLT 5

RESULT 8
PT0650
T-cell receptor beta chain V-D-J region (121-3BF) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0650
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regi
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0650
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:56:46 ; Search time 11.6667 Seconds
(without alignments)
57.715 Million cell updates/sec

Title: US-09-712-819C-5
Perfect score: 35
Sequence: 1 TDFTLTI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 457

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	13	37.1	6	2	PD0028	pev-kinin 2 - pena
2	12	34.3	5	2	A32516	cholecystokinin-5
3	12	34.3	5	2	PT0644	T-cell receptor be
4	12	34.3	7	2	B39127	phosphotransferase
5	12	34.3	7	2	PT0665	T-cell receptor be
6	11	31.4	4	2	I40697	biotin A - Citroba
7	11	31.4	7	2	E30608	Ig kappa chain V-I
8	10	28.6	6	2	PT0650	T-cell receptor be
9	10	28.6	7	2	S09066	globulin IV alpha
10	9	25.7	3	3	T13892	cytochrome-c oxida
11	9	25.7	5	2	E42364	flagellar protein
12	9	25.7	5	2	E60274	major protein anti
13	9	25.7	5	2	S68326	blood cell protein
14	9	25.7	5	2	S69237	surface protein te
15	9	25.7	6	2	A60986	N-formyl oligopept
16	9	25.7	6	2	B44510	hypothetical prote
17	9	25.7	6	2	A43766	28K ubiquitin-immu
18	9	25.7	6	2	I37263	Y protein - human
19	9	25.7	6	2	I65546	MHC H2-L antigen -
20	9	25.7	6	2	PT0587	T-cell receptor be
21	9	25.7	6	2	S29881	Na+/K+-exchanging
22	9	25.7	7	2	S25266	pilE protein - Esc
23	9	25.7	7	2	A25269	sex pheromone cAM3
24	9	25.7	7	2	A30812	sex pheromone cCF1
25	9	25.7	7	2	PT0611	T-cell receptor be
26	8	22.9	4	2	PT0697	T-cell receptor be
27	8	22.9	5	2	I39964	ribosomal protein
28	8	22.9	5	2	I39966	ribosomal protein
29	8	22.9	5	2	I39965	ribosomal protein
30	8	22.9	5	2	A44692	fulicin - giant Af
31	8	22.9	5	2	PT0729	T-cell receptor be
32	8	22.9	5	2	PT0590	T-cell receptor be
33	8	22.9	5	2	G44817	27.5 kda structura
34	8	22.9	5	2	I44817	27.5K structural p
35	8	22.9	5	2	E44817	27.5K structural p
36	8	22.9	5	2	C44817	28.5K structural p
37	8	22.9	5	2	A44817	28K structural pro
38	8	22.9	5	3	JT0870	phytosulfokine alp
39	8	22.9	6	2	A19780	transferrin - bovi
40	8	22.9	6	2	A46474	Fc epsilon RIIb -
41	8	22.9	6	2	PT0637	T-cell receptor be
42	8	22.9	6	2	PT0641	T-cell receptor be
43	8	22.9	7	2	E61491	seed protein ws-5
44	8	22.9	7	2	PS0254	18K protein 5507 -
45	8	22.9	7	2	PT0642	T-cell receptor be

ALIGNMENTS

RESULT 1
PD0028
pev-kinin 2 - penaeid shrimp (Penaeus vannamei) (fragment)
C;Species: Penaeus vannamei
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 19-May-2000
C;Accession: PD0028
R;Nieto, J.; Veelaert, D.; Derua, R.; Waelkens, E.; Cerstiaens, A.; Coast, G.; Devr
Biochem. Biophys. Res. Commun. 248, 406-411, 1998
A;Title: Identification of one tachykinin- and two kinin-related peptides in the br
A;Reference number: PD0027; MUID:98342103; PMID:9675150
A;Accession: PD0028
A;Molecule type: protein
A;Residues: 1-6 <NIE>
C;Comment: This peptide belongs to myotropic neuropeptides.

Query Match 37.1%; Score 13; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 DFT 4
              ||:
Db          1 DFS 3
```

RESULT 2
A32516
cholecystokinin-5 - dog
N;Alternate names: CCK-5
C;Species: Canis lupus familiaris (dog)
C;Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 18-Aug-2000
C;Accession: A32516
R;Shively, J.; Reeve Jr., J.R.; Eysselein, V.E.; Ben-Avram, C.; Vigna, S.R.; Walsh,
Am. J. Physiol. 252, G272-G275, 1987
A;Title: CCK-5: sequence analysis of a small cholecystokinin from canine brain and
A;Reference number: A32516; MUID:87153871; PMID:3826354
A;Accession: A32516
A;Molecule type: protein
A;Residues: 1-5 <SHI>
C;Comment: This peptide corresponds to the five carboxyl-terminal residues of chole
C;Superfamily: gastrin
C;Keywords: amidated carboxyl end; neuropeptide
F;5/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 34.3%; Score 12; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 DF 3
              ||
Db          4 DF 5
```

```
;       OPERATING SYSTEM:  PC/DOS/MS/DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/478,386A
;       FILING DATE:  07/JUN/1995
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Figg, E. Anthony
;       REGISTRATION NUMBER:  27,195
;       REFERENCE/DOCKET NUMBER:  2054-114A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 783-6040
;       TELEFAX:  (202) 783-6031
;   INFORMATION FOR SEQ ID NO:  46:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-478-386A-46

  Query Match             50.0%;  Score 16;  DB 2;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTIS 6
              ||:|
Db          3 LTVS 6


RESULT 14
US-08-292-597-46
; Sequence 46, Application US/08292597
; Patent No. 5834266
;  GENERAL INFORMATION:
;    APPLICANT:  Gerald R. Crabtree
;    APPLICANT:  Schreiber, Stuart L.
;    APPLICANT:  Spencer, David M.
;    APPLICANT:  Wandless, Thomas J.
;    APPLICANT:  Belshaw, Peter
;    TITLE OF INVENTION:  Regulated Apoptosis
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARIAD Pharmaceuticals, Inc.
;      STREET:  26 Landsdowne Street
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02139
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC/DOS/MS/DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/292,597
;      FILING DATE:  18/AUG/1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Figg, E. Anthony
;      REGISTRATION NUMBER:  27,195
;      REFERENCE/DOCKET NUMBER:  2054-108A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 783-6040
;      TELEFAX:  (202) 783-6031
;  INFORMATION FOR SEQ ID NO:  46:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-292-597-46

  Query Match             50.0%;  Score 16;  DB 2;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTIS 6
              ||:|
Db          3 LTVS 6


RESULT 15
US-08-388-653-46
; Sequence 46, Application US/08388653
; Patent No. 5869337
;  GENERAL INFORMATION:
;    APPLICANT:  Crabtree, Gerald R.
;    APPLICANT:  Schreiber, Stuart L.
;    APPLICANT:  Spencer, David M.
;    APPLICANT:  Wandless, Thomas J.
;    APPLICANT:  Belshaw, Peter
;    TITLE OF INVENTION:  REGULATED TRANSCRIPTION OF TARGETED
;    TITLE OF INVENTION:  GENES AND OTHER BIOLOGICAL EVENTS
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARIAD Pharmaceuticals, Inc.
;      STREET:  26 Landsdowne Street
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02139
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC/DOS/MS/DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/388,653
;      FILING DATE:  14-FEB-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/478,386
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Figg, E. Anthony
;      REGISTRATION NUMBER:  27,195
;      REFERENCE/DOCKET NUMBER:  2054-114A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 783-6040
;      TELEFAX:  (202) 783-6031
;  INFORMATION FOR SEQ ID NO:  46:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-388-653-46

  Query Match             50.0%;  Score 16;  DB 2;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTIS 6
              ||:|
Db          3 LTVS 6


Search completed: June  1, 2004, 11:20:17
Job time : 14.6667 secs
```

```
US-08-860-904-9
; Sequence 9, Application US/08860904
; Patent No. 6294654
; GENERAL INFORMATION:
;  APPLICANT: Sandlie, Inger
;  APPLICANT: Bogen, Bjarne
;  APPLICANT: Fossum, Sigbjorn
;  TITLE OF INVENTION: A Modified Immunoglobulin Molecule
;  TITLE OF INVENTION: Incorporating an Antigen in a No. 6294654-CDR
;  TITLE OF INVENTION: Loop Region
;  FILE REFERENCE: 9914-1
;  CURRENT APPLICATION NUMBER: US/08/860,904
;  CURRENT FILING DATE: 1997-09-29
;  EARLIER APPLICATION NUMBER: PCT/GB96/00116
;  EARLIER FILING DATE: 1996-01-19
;  EARLIER APPLICATION NUMBER: GB 9501079.9
;  EARLIER FILING DATE: 1995-01-19
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Mus sp.
US-08-860-904-9

  Query Match             50.0%;  Score 16;  DB 3;  Length 4;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TISS 7
              |:||
Db          1 TVSS 4


RESULT 11
US-09-301-593-47
; Sequence 47, Application US/09301593A
; Patent No. 6455677
; GENERAL INFORMATION:
;  APPLICANT: Park, John E.
;  APPLICANT: Garin-Chesa, Pilar
;  APPLICANT: Bamberger, Uwe
;  APPLICANT: Leger, Olivier
;  APPLICANT: Saldanha, Jose W.
;  APPLICANT: Rettig, Wolfgang J.
;  TITLE OF INVENTION: FAP-specific Antibody with Improved Producibility
;  FILE REFERENCE: 0652.1890001
;  CURRENT APPLICATION NUMBER: US/09/301,593A
;  CURRENT FILING DATE: 1999-04-29
;  EARLIER APPLICATION NUMBER: EP 98107925.4
;  EARLIER FILING DATE: 1998-04-30
;  EARLIER APPLICATION NUMBER: US 60/086,049
;  EARLIER FILING DATE: 1998-05-18
;  NUMBER OF SEQ ID NOS: 108
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-301-593-47

  Query Match             50.0%;  Score 16;  DB 4;  Length 4;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TISS 7
              |:||
Db          1 TVSS 4


RESULT 12
US-08-252-995D-7
; Sequence 7, Application US/08252995D
; Patent No. 5650501
;  GENERAL INFORMATION:
;    APPLICANT:  Dennis, James W
;    APPLICANT:  Heffernan, Mike
;    APPLICANT:  Fode, Carol
;    TITLE OF INVENTION:  NOVEL SERINE/THREONINE KINASE
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BERESKIN & PARR
;      STREET:  40 King Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/252,995D
;      FILING DATE:  02-JUN-1994
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kurdydyk, Linda M
;      REGISTRATION NUMBER:  34,971
;      REFERENCE/DOCKET NUMBER:  3153-96
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Mus musculus
US-08-252-995D-7

  Query Match             50.0%;  Score 16;  DB 1;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTISS 7
              ||:|:
Db          2 LTLSN 6


RESULT 13
US-08-478-386A-46
; Sequence 46, Application US/08478386A
; Patent No. 5830462
;  GENERAL INFORMATION:
;    APPLICANT:  Crabtree, Gerald R.
;    APPLICANT:  Schreiber, Stuart L.
;    APPLICANT:  Spencer, David M.
;    APPLICANT:  Wandless, Thomas J.
;    APPLICANT:  Belshaw, Peter
;    TITLE OF INVENTION:  REGULATED TRANSCRIPTION OF TARGETED
;    TITLE OF INVENTION:  GENES AND OTHER BIOLOGICAL EVENTS
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARIAD Pharmaceuticals, Inc.
;      STREET:  26 Landsdowne Street
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02139
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;  APPLICANT: Unilever PLC
;  TITLE OF INVENTION: New products comprising inactivated yeasts or moulds
;  TITLE OF INVENTION: provided with active antibodies
;  FILE REFERENCE: t-7055
;  CURRENT APPLICATION NUMBER: US/09/266,805
;  CURRENT FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: llama
US-09-266-805-5

  Query Match             53.1%;  Score 17;  DB 4;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTISS 7
              :|:||
Db          2 VTVSS 6


RESULT 7
US-09-530-139-14
; Sequence 14, Application US/09530139
; Patent No. 6670453
; GENERAL INFORMATION:
;  APPLICANT: FRENKEN, LEON GERARDUS
;  APPLICANT:  HOWELL, STEVEN
;  APPLICANT:  LEDEBOER, ADRIANUS MARINUS
;  APPLICANT:  VAN DER LOGT, CORNELIS PAUL
;  TITLE OF INVENTION: MULTIVALENT ANTIGEN-BINDING PROTEINS
;  FILE REFERENCE: 60113/268075/ASH
;  CURRENT APPLICATION NUMBER: US/09/530,139
;  CURRENT FILING DATE:  2000-04-27
;  PRIOR APPLICATION NUMBER: PCT/EP98/06991
;  PRIOR FILING DATE: 1998-10-27
;  PRIOR APPLICATION NUMBER: EP 97308538.4
;  PRIOR FILING DATE: 1997-10-27
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-09-530-139-14

  Query Match             53.1%;  Score 17;  DB 4;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTISS 7
              :|:||
Db          2 VTVSS 6


RESULT 8
US-08-739-401A-6
; Sequence 6, Application US/08739401A
; Patent No. 5837461
;  GENERAL INFORMATION:
;    APPLICANT:  Neitz, Maureen E.
;    APPLICANT:  Neitz, John F.
;    TITLE OF INVENTION:  DETECTION OF CONE-PHOTORECEPTOR-BASED
;    TITLE OF INVENTION:  VISION DISORDERS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  411 East Wisconsin Avenue
;      CITY:  Milwaukee
;      STATE:  Wisconsin
;      COUNTRY:  U.S.A.
;      ZIP:  53202-4497
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/739,401A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Baker, Jean C.
;      REGISTRATION NUMBER:  35,433
;      REFERENCE/DOCKET NUMBER:  650053.91151
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (414) 277-5709
;      TELEFAX:  (414) 271-3552
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-739-401A-6

  Query Match             53.1%;  Score 17;  DB 2;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLT 4
              ||:|
Db          2 FTVT 5


RESULT 9
US-08-753-750B-19
; Sequence 19, Application US/08753750B
; Patent No. 6610506
; GENERAL INFORMATION:
;  APPLICANT: Lo, Reggie Y.C.
;  APPLICANT:  Schryvers, Anthony B.
;  APPLICANT:  Potter, Andrew A.
;  TITLE OF INVENTION: TRANSFERRIN BINDING PROTEINS OF
;  TITLE OF INVENTION:  PASTEURELLA HAEMOLYTICA AND VACCINES CONTAINING THE SAME
;  FILE REFERENCE: A34762 021645.0105
;  CURRENT APPLICATION NUMBER: US/08/753,750B
;  CURRENT FILING DATE:  1996-11-29
;  PRIOR APPLICATION NUMBER: CA 2,164,274
;  PRIOR FILING DATE: 1995-12-01
;  PRIOR APPLICATION NUMBER: 60/008,569
;  PRIOR FILING DATE: 1995-12-01
;  NUMBER OF SEQ ID NOS: 68
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-08-753-750B-19

  Query Match             53.1%;  Score 17;  DB 4;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 3e+05;
  Matches    2;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLTISS 7
              |:|:::
Db          2 TITVTA 7


RESULT 10
```

```
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Representative linear modulating agent based on
;   OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
;   OTHER INFORMATION: recognition sequence
US-09-187-859-637

  Query Match             56.2%;  Score 18;  DB 4;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 3e+05;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLTISS 7
              ||:  ||
Db          1 FTIDSSS 7


RESULT 3
US-09-839-542B-637
; Sequence 637, Application US/09839542B
; Patent No. 6569996
; GENERAL INFORMATION:
;  APPLICANT: Blaschuk, Orest W.
;  APPLICANT:  Symonds, James Matthew
;  APPLICANT:  Gour, Barbara J.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
;  TITLE OF INVENTION:  CADHERIN-MEDIATED FUNCTIONS
;  FILE REFERENCE: 100086.407D1
;  CURRENT APPLICATION NUMBER: US/09/839,542B
;  CURRENT FILING DATE:  2001-04-20
;  NUMBER OF SEQ ID NOS: 4052
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 637
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Representative linear modulating agent based on
;   OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
;   OTHER INFORMATION: recognition sequence
US-09-839-542B-637

  Query Match             56.2%;  Score 18;  DB 4;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 3e+05;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLTISS 7
              ||:  ||
Db          1 FTIDSSS 7


RESULT 4
US-08-757-177-16
; Sequence 16, Application US/08757177
; Patent No. 6071718
;  GENERAL INFORMATION:
;    APPLICANT:  MUKERJI, PRADIP
;    APPLICANT:  HARDS, ROBERT G.
;    APPLICANT:  THURMOND, JENNIFER M.
;    APPLICANT:  LEONARD, AMANDA EUN-YEONG
;    TITLE OF INVENTION:  METHODS OF PRODUCING A RECOMBINANT PROTEIN
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  One Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,177
;      FILING DATE:  27-NOV-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Becker, Cheryl L.
;      REGISTRATION NUMBER:  35,441
;      REFERENCE/DOCKET NUMBER:  6004.US.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  847-935-1729
;      TELEFAX:  848-938-2623
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-757-177-16

  Query Match             53.1%;  Score 17;  DB 3;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TISS 7
              ||||
Db          2 TISS 5


RESULT 5
US-09-155-613A-59
; Sequence 59, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
;  APPLICANT: Boulanger, Pierre
;  APPLICANT:  Hong, Saw See
;  APPLICANT:  Karayan, Lucie
;  TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
;  FILE REFERENCE: 032751-036
;  CURRENT APPLICATION NUMBER: US/09/155,613A
;  CURRENT FILING DATE:  1998-09-30
;  PRIOR APPLICATION NUMBER: PCT/FR98/00184
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: FR 97/01005
;  PRIOR FILING DATE: 1997-01-30
;  PRIOR APPLICATION NUMBER: FR 97/11166
;  PRIOR FILING DATE: 1997-09-09
;  NUMBER OF SEQ ID NOS: 98
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Phagotope
US-09-155-613A-59

  Query Match             53.1%;  Score 17;  DB 4;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLT 4
              :|||
Db          3 YTLT 6


RESULT 6
US-09-266-805-5
; Sequence 5, Application US/09266805
; Patent No. 6517829
; GENERAL INFORMATION:
;  APPLICANT: Unilever N.V.
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 11:13:51 ; Search time 13.6667 Seconds
(without alignments)
26.443 Million cell updates/sec

Title: US-09-712-819C-1
Perfect score: 32
Sequence: 1 FTLTISS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 57228

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	62.5	7	4	US-08-753-750B-20	Sequence 20, Appl
2	18	56.2	7	4	US-09-187-859-637	Sequence 637, App
3	18	56.2	7	4	US-09-839-542B-637	Sequence 637, App
4	17	53.1	5	3	US-08-757-177-16	Sequence 16, Appl
5	17	53.1	6	4	US-09-155-613A-59	Sequence 59, Appl
6	17	53.1	6	4	US-09-266-805-5	Sequence 5, Appli
7	17	53.1	6	4	US-09-530-139-14	Sequence 14, Appl
8	17	53.1	7	2	US-08-739-401A-6	Sequence 6, Appli
9	17	53.1	7	4	US-08-753-750B-19	Sequence 19, Appl
10	16	50.0	4	3	US-08-860-904-9	Sequence 9, Appli
11	16	50.0	4	4	US-09-301-593-47	Sequence 47, Appl
12	16	50.0	6	1	US-08-252-995D-7	Sequence 7, Appli
13	16	50.0	6	2	US-08-478-386A-46	Sequence 46, Appl
14	16	50.0	6	2	US-08-292-597-46	Sequence 46, Appl
15	16	50.0	6	2	US-08-388-653-46	Sequence 46, Appl
16	16	50.0	6	2	US-08-473-985-46	Sequence 46, Appl
17	16	50.0	6	2	US-08-834-108-7	Sequence 7, Appli
18	16	50.0	6	2	US-08-483-898-46	Sequence 46, Appl
19	16	50.0	6	3	US-09-087-716-46	Sequence 46, Appl
20	16	50.0	6	3	US-09-157-753-46	Sequence 46, Appl
21	16	50.0	6	3	US-09-157-230-46	Sequence 46, Appl
22	16	50.0	6	3	US-09-087-811-46	Sequence 46, Appl
23	16	50.0	6	3	US-09-156-855-46	Sequence 46, Appl
24	16	50.0	6	3	US-09-158-010-46	Sequence 46, Appl
25	16	50.0	6	3	US-09-087-647-46	Sequence 46, Appl
26	16	50.0	6	4	US-09-302-629-46	Sequence 46, Appl
27	16	50.0	7	1	US-08-136-743B-55	Sequence 55, Appl
28	16	50.0	7	2	US-08-177-109A-7	Sequence 7, Appli
29	16	50.0	7	2	US-08-687-706-7	Sequence 7, Appli
30	16	50.0	7	3	US-09-040-216-28	Sequence 28, Appl
31	15	46.9	5	1	US-08-136-743B-63	Sequence 63, Appl
32	15	46.9	5	2	US-08-667-001-22	Sequence 22, Appl
33	15	46.9	5	3	US-09-040-216-55	Sequence 55, Appl
34	15	46.9	5	3	US-08-591-632-17	Sequence 17, Appl
35	15	46.9	5	3	US-08-591-632-23	Sequence 23, Appl
36	15	46.9	5	3	US-08-591-632-26	Sequence 26, Appl
37	15	46.9	5	4	US-09-611-451-17	Sequence 17, Appl
38	15	46.9	5	4	US-09-611-451-23	Sequence 23, Appl
39	15	46.9	5	4	US-09-611-451-26	Sequence 26, Appl
40	15	46.9	6	1	US-08-136-743B-62	Sequence 62, Appl
41	15	46.9	6	1	US-08-297-731-4	Sequence 4, Appli
42	15	46.9	6	1	US-08-290-448A-41	Sequence 41, Appl
43	15	46.9	6	1	US-08-290-448A-41	Sequence 41, Appl
44	15	46.9	6	1	US-08-175-069A-41	Sequence 41, Appl
45	15	46.9	6	2	US-08-622-720A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-753-750B-20
; Sequence 20, Application US/08753750B
; Patent No. 6610506
; GENERAL INFORMATION:
; APPLICANT: Lo, Reggie Y.C.
; APPLICANT: Schryvers, Anthony B.
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: TRANSFERRIN BINDING PROTEINS OF
; TITLE OF INVENTION: PASTEURELLA HAEMOLYTICA AND VACCINES CONTAINING THE SAME
; FILE REFERENCE: A34762 021645.0105
; CURRENT APPLICATION NUMBER: US/08/753,750B
; CURRENT FILING DATE: 1996-11-29
; PRIOR APPLICATION NUMBER: CA 2,164,274
; PRIOR FILING DATE: 1995-12-01
; PRIOR APPLICATION NUMBER: 60/008,569
; PRIOR FILING DATE: 1995-12-01
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Escherichia coli
US-08-753-750B-20

Query Match 62.5%; Score 20; DB 4; Length 7;
Best Local Similarity 42.9%; Pred. No. 3e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTLTISS 7
              |||:: :
Db          1 FTLSVDA 7
```

RESULT 2
US-09-187-859-637
; Sequence 637, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 637
; LENGTH: 7

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          3 LTISS 7
              ||| |
Db          1 LTIES 5
```

RESULT 14
US-10-159-006-47
; Sequence 47, Application US/10159006
; Publication No. US20030143229A1
; GENERAL INFORMATION:
; APPLICANT: Park, John E.
; APPLICANT: Garin-Chesa, Pilar
; APPLICANT: Bamberger, Uwe
; APPLICANT: Leger, Olivier
; APPLICANT: Saldanha, Jose W.
; APPLICANT: Rettig, Wolfgang J.
; TITLE OF INVENTION: FAPa-specific Antibody with Improved Producibility
; FILE REFERENCE: 0652.1890002
; CURRENT APPLICATION NUMBER: US/10/159,006
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/301,593
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: EP 98107925.4
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: US 60/086,049
; PRIOR FILING DATE: 1998-05-18
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-159-006-47

Query Match 50.0%; Score 16; DB 14; Length 4;
Best Local Similarity 75.0%; Pred. No. 1e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          4 TISS 7
              |:||
Db          1 TVSS 4
```

RESULT 15
US-10-020-354-102
; Sequence 102, Application US/10020354
; Publication No. US20030190311A1
; GENERAL INFORMATION:
; APPLICANT: DALL'ACQUA, WILLIAM
; APPLICANT: JOHNSON, LESLIE
; APPLICANT: WARD, ELIZABETH SALLY
; TITLE OF INVENTION: MOLECULES WITH EXTENDED HALF-LIVES, COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: 10271-027
; CURRENT APPLICATION NUMBER: US/10/020,354
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/254,884
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/238,760
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 102
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-020-354-102

Query Match 50.0%; Score 16; DB 14; Length 6;
Best Local Similarity 60.0%; Pred. No. 1e+06;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 TLTIS 6
              |::||
Db          2 TISIS 6
```

Search completed: June 1, 2004, 11:31:22
Job time : 35.6667 secs

Query Match 53.1%; Score 17; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 LTIS 6
              ||||
Db          2 LTIS 5
```

RESULT 10
US-09-530-139-14
; Sequence 14, Application US/09530139
; Publication No. US20030092892A1
; GENERAL INFORMATION:
; APPLICANT: FRENKEN, LEON GERARDUS
; APPLICANT: HOWELL, STEVEN
; APPLICANT: LEDEBOER, ADRIANUS MARINUS
; APPLICANT: VAN DER LOGT, CORNELIS PAUL
; TITLE OF INVENTION: MULTIVALENT ANTIGEN-BINDING PROTEINS
; FILE REFERENCE: 60113/268075/ASH
; CURRENT APPLICATION NUMBER: US/09/530,139
; CURRENT FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/EP98/06991
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: EP 97308538.4
; PRIOR FILING DATE: 1997-10-27
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-530-139-14

Query Match 53.1%; Score 17; DB 10; Length 6;
Best Local Similarity 60.0%; Pred. No. 1e+06;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 LTISS 7
              :|:||
Db          2 VTVSS 6
```

RESULT 11
US-10-698-489-17
; Sequence 17, Application US/10698489
; Publication No. US20040067523A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001103CON
; CURRENT APPLICATION NUMBER: US/10/698,489
; CURRENT FILING DATE: 2003-11-03
; PRIOR APPLICATION NUMBER: To Be Assigned
; PRIOR FILING DATE: 2003-11-03
; PRIOR APPLICATION NUMBER: 09/777,921
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-698-489-17

Query Match 53.1%; Score 17; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 LTIS 6
              ||||
Db          2 LTIS 5
```

RESULT 12
US-10-156-820-59
; Sequence 59, Application US/10156820
; Publication No. US20020150558A1
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/10/156,820
; CURRENT FILING DATE: 2002-06-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotope
US-10-156-820-59

Query Match 53.1%; Score 17; DB 13; Length 6;
Best Local Similarity 75.0%; Pred. No. 1e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FTLT 4
              :|||
Db          3 YTLT 6
```

RESULT 13
US-10-097-175-97
; Sequence 97, Application US/10097175
; Publication No. US20030045680A1
; GENERAL INFORMATION:
; APPLICANT: JOYAL, JOHN L.
; APPLICANT: MUELLER, JOHN
; APPLICANT: OZA, VIBHA B.
; APPLICANT: FINDEIS, MARK A.
; TITLE OF INVENTION: PEPTIDIC MODULATORS OF THE ANDROGEN RECEPTOR
; FILE REFERENCE: PPI-110
; CURRENT APPLICATION NUMBER: US/10/097,175
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,240
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/352,399
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Androgen Receptor Binding Polypeptides
US-10-097-175-97

Query Match 53.1%; Score 17; DB 14; Length 6;
Best Local Similarity 80.0%; Pred. No. 1e+06;

```
  Query Match           56.2%;  Score 18;  DB 15;  Length 7;
  Best Local Similarity  57.1%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLTISS 7
              ||:  ||
Db          1 FTIDSSS 7


RESULT 6
US-09-788-006-107
; Sequence 107, Application US/09788006
; Publication No. US20030036093A1
; GENERAL INFORMATION:
;  APPLICANT: Floudas, Christopher A.
;  APPLICANT:  Klepeis, John L.
;  TITLE OF INVENTION: Methods of Ab Initio Prediction of Alpha Helices, Beta Sheets,and
;  TITLE OF INVENTION:  Polypeptide Tertiary Structures
;  FILE REFERENCE: PU-0007
;  CURRENT APPLICATION NUMBER: US/09/788,006
;  CURRENT FILING DATE:  2001-02-16
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 107
;   LENGTH: 5
;   TYPE: PRT
;   ORGANISM: Streptomyces griseus
US-09-788-006-107

  Query Match           53.1%;  Score 17;  DB 10;  Length 5;
  Best Local Similarity  75.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLT 4
              ||:|
Db          2 FTVT 5


RESULT 7
US-09-788-006-108
; Sequence 108, Application US/09788006
; Publication No. US20030036093A1
; GENERAL INFORMATION:
;  APPLICANT: Floudas, Christopher A.
;  APPLICANT:  Klepeis, John L.
;  TITLE OF INVENTION: Methods of Ab Initio Prediction of Alpha Helices, Beta Sheets,and
;  TITLE OF INVENTION:  Polypeptide Tertiary Structures
;  FILE REFERENCE: PU-0007
;  CURRENT APPLICATION NUMBER: US/09/788,006
;  CURRENT FILING DATE:  2001-02-16
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
;   LENGTH: 5
;   TYPE: PRT
;   ORGANISM: Streptomyces griseus
US-09-788-006-108

  Query Match           53.1%;  Score 17;  DB 10;  Length 5;
  Best Local Similarity  75.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLT 4
              ||:|
Db          1 FTVT 4


RESULT 8
US-10-154-971-29
; Sequence 29, Application US/10154971
; Publication No. US20030088074A1
;   GENERAL INFORMATION:
;        APPLICANT: Hamers, Raymond
;                   Muyldermans, Serge
;        TITLE OF INVENTION: VARIABLE FRAGMENTS OF IMMUNOGLOBULINS -
;                            USE FOR THERAPEUTIC OR VETERINARY PURPOSES
;        NUMBER OF SEQUENCES: 29
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SPENCER & FRANK
;             STREET: 1100 New York Avenue, N.W., Suite 300 East
;             CITY: Washington
;             STATE: DC
;             COUNTRY: USA
;             ZIP: 20005
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/154,971
;             FILING DATE: 28-May-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/945,244
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: PCT/EP96/01725
;             FILING DATE: 25-APR-1996
;             APPLICATION NUMBER: EP 95400932.0
;             FILING DATE: 25-APR-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Gollin, Michael A.
;             REGISTRATION NUMBER: 31,957
;             REFERENCE/DOCKET NUMBER: GUPLA 0003
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 202-414-4000
;             TELEFAX: 202-414-4040
;   INFORMATION FOR SEQ ID NO: 29:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 5 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-10-154-971-29

  Query Match           53.1%;  Score 17;  DB 14;  Length 5;
  Best Local Similarity  60.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTISS 7
              :|:||
Db          1 VTVSS 5


RESULT 9
US-09-777-921A-17
; Sequence 17, Application US/09777921A
; Patent No. US20020115136A1
; GENERAL INFORMATION:
;  APPLICANT: MERKULOV et al.
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL001103
;  CURRENT APPLICATION NUMBER: US/09/777,921A
;  CURRENT FILING DATE:  2002-02-07
;  NUMBER OF SEQ ID NOS: 126
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-777-921A-17
```

```
RESULT 2
US-10-267-565-11
; Sequence 11, Application US/10267565
; Publication No. US20030204059A1
; GENERAL INFORMATION:
;  APPLICANT: Gately, Maurice
;  APPLICANT:  Gubler, Ulrich
;  APPLICANT:  Hulmes, Jeffery
;  APPLICANT:  Podlaski, Frank
;  APPLICANT:  Stern, Alvin
;  TITLE OF INVENTION: PURIFICATION AND CHARACTERIZATION OF CYTOTOXIC
;  TITLE OF INVENTION:  LYMPHOCYTE MATURATION FACTOR AND MONOCLONAL
;  TITLE OF INVENTION:  ANTIBODIES THERETO
;  FILE REFERENCE: 11126-005
;  CURRENT APPLICATION NUMBER: US/10/267,565
;  CURRENT FILING DATE:  2003-01-23
;  PRIOR APPLICATION NUMBER: 09/401,839
;  PRIOR FILING DATE: 1999-09-22
;  PRIOR APPLICATION NUMBER: 08/459,151
;  PRIOR FILING DATE: 1995-06-02
;  PRIOR APPLICATION NUMBER: 08/205,011
;  PRIOR FILING DATE: 1994-03-02
;  PRIOR APPLICATION NUMBER: 07/857,023
;  PRIOR FILING DATE: 1992-03-24
;  PRIOR APPLICATION NUMBER: 07/572,284
;  PRIOR FILING DATE: 1990-08-27
;  PRIOR APPLICATION NUMBER: 07/520,935
;  PRIOR FILING DATE: 1990-05-09
;  PRIOR APPLICATION NUMBER: 07/455,708
;  PRIOR FILING DATE: 1989-12-22
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-267-565-11

  Query Match             56.2%;  Score 18;  DB 12;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLTI 5
              ||||
Db          1 TLTI 4


RESULT 3
US-10-006-869-637
; Sequence 637, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
;  APPLICANT: Blaschuk, Orest W.
;  APPLICANT:  Symonds, James Matthew
;  APPLICANT:  Gour, Barbara J.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
;  TITLE OF INVENTION:  CADHERIN-MEDIATED FUNCTIONS
;  FILE REFERENCE: 100086.407C7
;  CURRENT APPLICATION NUMBER: US/10/006,869
;  CURRENT FILING DATE:  2001-12-03
;  NUMBER OF SEQ ID NOS: 4052
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 637
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Representative linear modulating agent based on
;   OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
;   OTHER INFORMATION: recognition sequence
US-10-006-869-637

  Query Match             56.2%;  Score 18;  DB 14;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLTISS 7
              ||:  ||
Db          1 FTIDSSS 7


RESULT 4
US-10-349-507-11
; Sequence 11, Application US/10349507
; Publication No. US20030199002A1
; GENERAL INFORMATION:
;  APPLICANT: Hekimi , Siegfried
;  APPLICANT:  Jiang , Ning
;  APPLICANT:  Benard, Claire
;  APPLICANT:  Kebir , Hania
;  APPLICANT:  McCright, Brenton
;  APPLICANT:  Lakowski , Bernard
;  TITLE OF INVENTION: CLK-2 NUCLEIC ACIDS, POLYPEPTIDES, AND USES THEREOF
;  FILE REFERENCE: 11202-006-999
;  CURRENT APPLICATION NUMBER: US/10/349,507
;  CURRENT FILING DATE:  2003-01-22
;  PRIOR APPLICATION NUMBER: 10/312,187
;  PRIOR FILING DATE: 2002-12-20
;  PRIOR APPLICATION NUMBER: PCT/CA01/00913
;  PRIOR FILING DATE: 2001-06-20
;  PRIOR APPLICATION NUMBER: 60/213,174
;  PRIOR FILING DATE: 2000-06-22
;  PRIOR APPLICATION NUMBER: 60/254,932
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-349-507-11

  Query Match             56.2%;  Score 18;  DB 14;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLTI 5
              ||||
Db          1 TLTI 4


RESULT 5
US-10-395-032-637
; Sequence 637, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
;  APPLICANT: Blaschuk, Orest W.
;  APPLICANT:  Symonds, James Matthew
;  APPLICANT:  Gour, Barbara J.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
;  TITLE OF INVENTION:  CADHERIN-MEDIATED FUNCTIONS
;  FILE REFERENCE: 100086.407C9
;  CURRENT APPLICATION NUMBER: US/10/395,032
;  CURRENT FILING DATE:  2003-03-21
;  NUMBER OF SEQ ID NOS: 4052
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 637
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Representative linear modulating agent based on
;   OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
;   OTHER INFORMATION: recognition sequence
US-10-395-032-637
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 11:18:42 ; Search time 34.6667 Seconds
(without alignments)
56.387 Million cell updates/sec

Title: US-09-712-819C-1
Perfect score: 32
Sequence: 1 FTLTISS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 49082

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	22	68.8	6	14	US-10-150-654A-18	Sequence 18, Appl
2	18	56.2	7	12	US-10-267-565-11	Sequence 11, Appl
3	18	56.2	7	14	US-10-006-869-637	Sequence 637, App
4	18	56.2	7	14	US-10-349-507-11	Sequence 11, Appl
5	18	56.2	7	15	US-10-395-032-637	Sequence 637, App
6	17	53.1	5	10	US-09-788-006-107	Sequence 107, App
7	17	53.1	5	10	US-09-788-006-108	Sequence 108, App
8	17	53.1	5	14	US-10-154-971-29	Sequence 29, Appl
9	17	53.1	6	9	US-09-777-921A-17	Sequence 17, Appl
10	17	53.1	6	10	US-09-530-139-14	Sequence 14, Appl
11	17	53.1	6	12	US-10-698-489-17	Sequence 17, Appl
12	17	53.1	6	13	US-10-156-820-59	Sequence 59, Appl
13	17	53.1	6	14	US-10-097-175-97	Sequence 97, Appl
14	16	50.0	4	14	US-10-159-006-47	Sequence 47, Appl
15	16	50.0	6	14	US-10-020-354-102	Sequence 102, App
16	16	50.0	6	16	US-10-054-712-46	Sequence 46, Appl
17	16	50.0	7	9	US-09-832-723-69	Sequence 69, Appl
18	16	50.0	7	9	US-09-996-288-164	Sequence 164, App
19	16	50.0	7	10	US-09-996-265-164	Sequence 164, App
20	16	50.0	7	14	US-10-303-331-69	Sequence 69, Appl
21	16	50.0	7	14	US-10-022-066-214	Sequence 214, App
22	16	50.0	7	15	US-10-461-863-164	Sequence 164, App
23	15	46.9	5	14	US-10-286-186-3	Sequence 3, Appli
24	15	46.9	5	14	US-10-286-186-4	Sequence 4, Appli
25	15	46.9	5	14	US-10-315-964A-394	Sequence 394, App
26	15	46.9	5	14	US-10-317-251A-394	Sequence 394, App
27	15	46.9	5	14	US-10-317-252A-394	Sequence 394, App
28	15	46.9	6	9	US-09-876-388-6	Sequence 6, Appli
29	15	46.9	6	14	US-10-105-930-39	Sequence 39, Appl
30	15	46.9	6	14	US-10-006-869-650	Sequence 650, App
31	15	46.9	6	14	US-10-287-892-6	Sequence 6, Appli
32	15	46.9	6	14	US-10-288-340-6	Sequence 6, Appli
33	15	46.9	6	14	US-10-315-964A-402	Sequence 402, App
34	15	46.9	6	14	US-10-317-251A-402	Sequence 402, App
35	15	46.9	6	14	US-10-317-252A-402	Sequence 402, App
36	15	46.9	6	15	US-10-395-032-650	Sequence 650, App
37	15	46.9	6	15	US-10-394-980-243	Sequence 243, App
38	15	46.9	6	16	US-10-418-972-55	Sequence 55, Appl
39	15	46.9	6	16	US-10-374-466-47	Sequence 47, Appl
40	15	46.9	7	9	US-09-056-160B-5	Sequence 5, Appli
41	15	46.9	7	9	US-09-056-160B-124	Sequence 124, App
42	15	46.9	7	9	US-09-876-388-7	Sequence 7, Appli
43	15	46.9	7	9	US-09-734-417-12	Sequence 12, Appl
44	15	46.9	7	9	US-09-095-881-8	Sequence 8, Appli
45	15	46.9	7	9	US-09-947-137-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-10-150-654A-18
; Sequence 18, Application US/10150654A
; Publication No. US20030198595A1
; GENERAL INFORMATION:
; APPLICANT: GOLDENBERG, DAVID M.
; APPLICANT: HANSEN, HANS J.
; APPLICANT: LEUNG, SHUI-ON
; APPLICANT: MCBRIDE, WILLIAM J.
; APPLICANT: QU, ZHENGXING
; TITLE OF INVENTION: PRODUCTION AND USE OF NOVEL PEPTIDE-BASED AGENTS FOR
; TITLE OF INVENTION: USE WITH BI-SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733-1085
; CURRENT APPLICATION NUMBER: US/10/150,654A
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: 09/382,186
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: 09/823,746
; PRIOR FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-150-654A-18

Query Match 68.8%; Score 22; DB 14; Length 6;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 TLTISS 7
              |:|:||
Db          1 TVTVSS 6
```

CC amino acid residues of the first segment, together with an additional
CC amino acid residue(s), adjacent to the first amino acid sequence. The
CC beta-sheets and disulphide bridges of a polypeptide are determined by
CC identifying a superstructure encompassing alternative beta-strand
CC arrangements, representing the superstructure by an integer-linear
CC programming mathematical model, with the model maximising the total
CC hydrophobic contact energy, and then solving the model. The method can
CC also apply energy modelling with a full atom force field potential to
CC generate a constrained global optimisation problem, optionally applying a
CC torsion angle dynamics algorithm and then solving the constrained global
CC optimisation problem to determine the three-dimensional structure of the
CC polypeptide. The method is useful for determining the tertiary structure
CC of a polypeptide and is accurate and reliable. The sequences presented in
CC ABU12184-ABU122235 are the pentapeptides of the Streptomyces griseus
CC immunoglobulin-binding domain from streptococcal protein G, 1GB1, which
CC were used to predict the alpha-helical regions in an example of the
CC method of the invention
XX
SQ Sequence 5 AA;

Query Match 53.1%; Score 17; DB 6; Length 5;
Best Local Similarity 75.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FTLT 4
              ||:|
Db          2 FTVT 5
```

Search completed: June 1, 2004, 11:16:14
Job time : 49.6667 secs

CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC mammal, enhancing delivery of a drug through the skin of a mammal,
CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC expressing cell, preventing or treating obesity in a mammal, stimulating
CC blood vessel regression in a mammal, enhancing drug delivery to the
CC central nervous system, treating a demyelinating neurological disease,
CC increasing vasopermeability in a mammal, enhancing adhesion of
CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC a mammal, or preventing pregnancy in a mammal. They can also be used for
CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC -related macular degeneration, multiple sclerosis and diabetes. The
CC products can also be used for detection and diagnosis and in bioreactors.
CC AAY60592 to AAY64572 represent specifically claimed peptides, and
CC AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 56.2%; Score 18; DB 3; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 FTLTISS 7
              ||:  ||
Db          1 FTIDSSS 7
```

RESULT 14
ABB84102
ID ABB84102 standard; protein; 7 AA.
XX
AC ABB84102;
XX
DT 09-SEP-2002 (first entry)
XX
DE Human single chain diabody CEAGal fragment #2.
XX
KW Antibody; multimer; cytostatic; detection; immunoassay; tumour.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 6
FT /note= "Unidentified residue"
XX
PN DE10060140-A1.
XX
PD 06-JUN-2002.
XX
PF 04-DEC-2000; 2000DE-01060140.
XX
PR 04-DEC-2000; 2000DE-01060140.
XX
PA (VECT-) VECTRON THERAPEUTICS IMT AG.
XX
PI Kontermann R;
XX
DR WPI; 2002-520984/56.
XX
PT New multimer of single-chain antibodies, useful e.g. for diagnosis or
PT drug delivery, has four variable chains linked through peptides of
PT controlled length.
XX
PS Example 1; Fig 2B; 16pp; German.
XX
CC This invention describes a novel multimer containing at least two
CC antibodies, each, independently, comprising a polypeptide chain of
CC structure V1-PA-V2-PM-V3-PB-V4 where V1-V4 = variable domains; PA, PB and
CC PM = peptide linkers with PA and PB containing 0 or 1 amino acid. The
CC products of the invention have cytostatic activity. The multimers
CC described can be used for diagnostic detection, in essentially standard
CC immunoassay methods, and for binding at least one component and/or fused
CC protein or peptide to a cell, especially to deliver a toxin or for
CC infection, transformation or transfection of the cell. (I), or nucleic
CC acid (II) encoding it, or vectors containing (II), are useful for
CC treating tumours. The multimers are relatively small (about 110 kDa for a
CC dimer) but have multiple binding sites to ensure high binding affinity.
CC This sequence represents a fragment of the single chain diabody CEAGal
CC (composed of a tumour marker carcinoembryonic antigen fragment and E.
CC coli beta-galactosidase fragment) described in the method of the
CC invention
XX
SQ Sequence 7 AA;

Query Match 56.2%; Score 18; DB 5; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 TLTIS 6
              |:|:|
Db          1 TVTVS 5
```

RESULT 15
ABU12234
ID ABU12234 standard; peptide; 5 AA.
XX
AC ABU12234;
XX
DT 18-FEB-2003 (first entry)
XX
DE Streptococcal protein G, 1GB1, alpha helix prediction pentapeptide #51.
XX
KW ASTRO-FOLD; ab anitio; alpha-helix; beta-sheet; disulphide bridge;
KW atomistic modelling; low energy conformation; entropic energy;
KW free energy; equilibrium probability; helical cluster; integer-linear;
KW mathematical model; hydrophobic contact energy;
KW atom force field potential; global optimisation; torsion angle;
KW dynamic algorithm; three-dimensional structure; tertiary structure;
KW immunoglobulin-binding domain; protein G; 1GB1.
XX
OS Streptomyces griseus.
XX
PN WO200279872-A2.
XX
PD 10-OCT-2002.
XX
PF 19-FEB-2002; 2002WO-US004644.
XX
PR 16-FEB-2001; 2001US-00788006.
XX
PA (UYPR-) UNIV PRINCETON.
XX
PI Floudas CA, Klepeis JL;
XX
DR WPI; 2003-058449/05.
XX
PT Determining helix regions, and beta sheets of polypeptide, involves
PT partitioning peptide, atomistic modeling by selected force, generating
PT ensemble low energy, and calculating free energies for each peptide.
XX
PS Example 1; Page 59; 147pp; English.
XX
CC The invention discloses a ASTRO-FOLD approach for the ab anitio
CC prediction method for determining the existence and location of alpha-
CC helix regions and arrangement of beta-sheets and disulphide bridges of a
CC polypeptide. The method comprises defining the first segment of the amino
CC acid sequence, performing atomistic modelling upon each segment,
CC generating an ensemble of low energy conformations, determining the
CC entropic and free energy for each segment and then ascertaining the
CC equilibrium probabilities for helical clusters. The segments consist of
CC pentapeptides, with each further segment including a majority of the

XX
PI Lai T;
XX
DR WPI; 2003-421400/39.
XX
PT New phospholipase C-zeta gene, useful for preparing a composition for
PT diagnosing or treating infertility.
XX
PS Claim 10; Page 68; 107pp; English.
XX
CC ABR55412-20 represent conserved region of a phospholipase C (PLC)-zeta
CC polypeptide. PLC-zeta polypeptides are capable of triggering calcium
CC oscillations in oocytes. The PLC-zeta isoform is expressed specifically
CC in mammalian sperm, and is an essential protein for mammalian
CC fertilisation and embryo development. PLC-zeta polynucleotides are useful
CC for preparing a composition for diagnosing or treating infertility
XX
SQ Sequence 5 AA;

Query Match 56.2%; Score 18; DB 6; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLTI 5
||||
Db 1 TLTI 4

RESULT 12
AAW58711
ID AAW58711 standard; peptide; 7 AA.
XX
AC AAW58711;
XX
DT 17-SEP-1998 (first entry)
XX
DE Tryptic 40 kD subunit CLMF peptide off PVDF #4.
XX
KW Cytotoxic lymphocyte maturation factor; CLMF; cytokine; synergize;
KW interleukin 2; IL-2; human; lymphoblastoid cell; detection; purifying;
KW proliferation; cytotoxic T cell; transplantation; antibody.
XX
OS Homo sapiens.
XX
PN US5780597-A.
XX
PD 14-JUL-1998.
XX
PF 02-JUN-1995; 95US-00460061.
XX
PR 22-DEC-1989; 89US-00455708.
PR 09-MAY-1990; 90US-00520935.
PR 27-AUG-1990; 90US-00572284.
PR 24-MAR-1992; 92US-00857023.
PR 02-MAR-1994; 94US-00205011.
XX
PA (HOFF) HOFFMANN LA ROCHE INC.
XX
PI Podlaski FJ, Stern AS, Gately MK, Pan YE, Hulmes JD;
PI Chizzonite RA, Gubler UA;
XX
DR WPI; 1998-413150/35.
XX
PT New antibodies to cytotoxic lymphocyte maturation factor - useful for
PT detecting, purifying, and/or blocking proliferation and activation of
PT cytotoxic T cells, such as in transplantation(s).
XX
PS Example 3; Col 23; 71pp; English.
XX
CC An isolated antibody has been developed which binds specifically to
CC cytotoxic lymphocyte maturation factor (CLMF) and neutralises CLMF
CC induced proliferation by more than 50% in a CLMF dependent T cell growth
CC assay and/or inhibits binding of at least 60% of the factor to
CC phytohaemagglutinin (PHA) activated peripheral blood lymphocyte (PBL)
CC blasts as determined in a CLMF receptor binding assay. CLMF is a
CC heterodimeric protein having a molecular weight band of 75 kD, determined
CC by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)
CC under non-reducing and/or reducing conditions, providing a first subunit
CC having a molecular weight band of 40 kD and a second subunit having a
CC molecular weight of 35 kD. The present sequence represents a tryptic 40
CC kD CLMF peptide off polyvinylidene difluoride (PVDF), from an example of
CC the present invention. The antibody can be used for the purification
CC and/or detection of CLMF. It is also used in therapeutic treatments which
CC require selective blocking of proliferation and activation of cytotoxic T
CC cells (CTLs) such as in transplantation
XX
SQ Sequence 7 AA;

Query Match 56.2%; Score 18; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLTI 5
||||
Db 1 TLTI 4

RESULT 13
AAY64225
ID AAY64225 standard; peptide; 7 AA.
XX
AC AAY64225;
XX
DT 02-MAR-2000 (first entry)
XX
DE Cadherin-related neuronal receptor CAR sequence SEQ ID NO:637.
XX
KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW inhibition; cadherin extracellular domain; cell adhesion recognition;
KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW neurological disease.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9957149-A2.
XX
PD 11-NOV-1999.
XX
PF 05-MAY-1999; 99WO-CA000363.
XX
PR 05-MAY-1998; 98US-00073040.
PR 06-NOV-1998; 98US-00187859.
PR 20-JAN-1999; 99US-00234395.
PR 08-MAR-1999; 99US-00264516.
XX
PA (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI Blaschuk OW, Gour BJ, Byers S;
XX
DR WPI; 2000-038791/03.
XX
PT New cadherin modulating agents, used for modulating nonclassical cadherin
PT -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS Claim 99; Page 216; 252pp; English.
XX
CC The present invention describes cadherin modulating agents (MA)
CC comprising peptides which comprise a nonclassical cadherin cell adhesion
CC recognition (CAR) sequence. The MAs can be used for modulating
CC nonclassical cadherin-mediated functions. They can be used for e.g.

```
  Query Match           65.6%;  Score 21;  DB 3;  Length 7;
  Best Local Similarity  66.7%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLTIS 6
              ::||||
Db          2 YSLTIS 7


RESULT 9
AAU85454
ID   AAU85454 standard; peptide; 5 AA.
XX
AC   AAU85454;
XX
DT   21-MAY-2002  (first entry)
XX
DE   Human colon specific polypeptide antibody binding site #31.
XX
KW   Human; colon specific gene; CSG; cytostatic; metastasis;
KW   colon cancer staging; antibody binding site.
XX
OS   Homo sapiens.
XX
PN   WO200206515-A2.
XX
PD   24-JAN-2002.
XX
PF   17-JUL-2001; 2001WO-US022454.
XX
PR   17-JUL-2000; 2000US-00618596.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Macina RA,  Sun Y;
XX
DR   WPI; 2002-171815/22.
XX
PT   Diagnosing, staging or monitoring colon cancer involves determining a
PT   colon specific gene in cells, tissues or body fluids in patient, and
PT   comparing it with levels of the gene from a normal human control.
XX
PS   Disclosure; Page 21; 52pp; English.
XX
CC   The invention relates to diagnosing the presence of colon cancer,
CC   metastases of colon cancer, staging colon cancer, monitoring colon cancer
CC   for the onset of metastasis or monitoring a change in stage of colon
CC   cancer in a patient. The method involves determining a colon specific
CC   gene (CSG) in cells, tissues or bodily fluids and comparing it with
CC   levels of CSG in cells, tissues or bodily fluids from a normal human
CC   control. Colon cancer can be treated by administering a molecule which
CC   down regulates the expression or activity of CSG. An immune response
CC   against a target cell expressing CSG can be induced by delivering an
CC   immunologically stimulatory amount of a CSG protein to a patient, so that
CC   an immune response is mounted. Therapeutic agents are useful for imaging
CC   colon cancer in a patient by administering an agent labelled with
CC   paramagnetic ions or a radioisotope to the patient. They are also useful
CC   for preventing the onset of colon cancer, and in diagnosis and treatment
CC   of the disease. Sequences AAU85424-AAU85502 represent human colon
CC   specific protein antibody binding sites used in the method of the
CC   invention
XX
SQ   Sequence 5 AA;

  Query Match           62.5%;  Score 20;  DB 5;  Length 5;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLT 4
              ||||
Db          2 FTLT 5


RESULT 10
AAM51422
ID   AAM51422 standard; peptide; 6 AA.
XX
AC   AAM51422;
XX
DT   08-JAN-2002  (first entry)
XX
DE   Integrin activating peptide SEQ ID NO 1.
XX
KW   Integrin stimulant; vulnerary; injury healing;
KW   postsurgical tissue recovery.
XX
OS   Unidentified.
XX
PN   JP2001213898-A.
XX
PD   07-AUG-2001.
XX
PF   31-JAN-2000; 2000JP-00022469.
XX
PR   31-JAN-2000; 2000JP-00022469.
XX
PA   (HISM ) HISAMITSU PHARM CO LTD.
XX
DR   WPI; 2001-629610/73.
XX
PT   An injury healing and postsurgical tissue recovering integrin activating
PT   peptide.
XX
PS   Claim 1; Page 3; 11pp; Japanese.
XX
CC   The invention relates to novel peptides with vulnerary activity, useful
CC   for injury healing and postsurgical tissue recovery by acting as an
CC   integrin stimulant
XX
SQ   Sequence 6 AA;

  Query Match           59.4%;  Score 19;  DB 4;  Length 6;
  Best Local Similarity  60.0%;  Pred. No. 1.4e+06;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLTI 5
              :|:||
Db          1 YTITI 5


RESULT 11
ABR55418
ID   ABR55418 standard; peptide; 5 AA.
XX
AC   ABR55418;
XX
DT   29-JUL-2003  (first entry)
XX
DE   Amino acid sequence of a conserved region of PLC-zeta protein.
XX
KW   Phospholipase C zeta; PLC-zeta; calcium oscillation; oocyte; sperm;
KW   embryo; infertility.
XX
OS   Unidentified.
XX
PN   WO2003035678-A2.
XX
PD   01-MAY-2003.
XX
PF   18-OCT-2002; 2002WO-GB004739.
XX
PR   24-OCT-2001; 2001GB-00025498.
PR   28-JUN-2002; 2002GB-00014945.
XX
PA   (UYWA-) UNIV WALES COLLEGE OF MEDICINE.
```

XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 2000-665002/64.
XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding fragments.
XX
PS Disclosure; Page 15; 68pp; English.
XX
CC The present invention is concerned with producing scaffold proteins based
CC upon the human CTLA-4 SCA domain. These scaffold proteins can be used as
CC a scaffold to bind antigen- or receptor-binding fragments. These can be
CC used in the treatment of diseases such as cancer, atherosclerosis,
CC thrombosis, osteoporosis, rheumatoid arthritis and diabetic retinopathy.
CC Sequences AAB29930-B29939 were used in the production of the proteins of
CC the invention
XX
SQ Sequence 7 AA;

Query Match 75.0%; Score 24; DB 3; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLTIS 6
||| :||
Db 2 FTLSIS 7

RESULT 7
AAY40735
ID AAY40735 standard; peptide; 7 AA.
XX
AC AAY40735;
XX
DT 01-DEC-1999 (first entry)
XX
DE S4 derivative #9, beta strand of scaffold protein structure.
XX
KW Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.
XX
OS Synthetic.
XX
PN EP947582-A1.
XX
PD 06-OCT-1999.
XX
PF 31-MAR-1998; 98EP-00870065.
XX
PR 31-MAR-1998; 98EP-00870065.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 1999-542958/46.
XX
PT New scaffold protein, useful for stabilizing antigens used as vaccines.
XX
PS Disclosure; Page 6; 105pp; English.
XX
CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides (AAY40601-
CC Y40609) together form a single-chain scaffold protein which contains at
CC least 1 disulfide bond, contains less than 10% alpha helix and contains
CC at least 6 beta-strands. The scaffold protein is constructed of beta
CC strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines
XX
SQ Sequence 7 AA;

Query Match 65.6%; Score 21; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLTIS 6
::||||
Db 2 YSLTIS 7

RESULT 8
AAB30073
ID AAB30073 standard; peptide; 7 AA.
XX
AC AAB30073;
XX
DT 09-FEB-2001 (first entry)
XX
DE Scaffold protein SCA S4 peptide SEQ ID NO: 134.
XX
KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.
XX
OS Synthetic.
XX
PN WO200060070-A1.
XX
PD 12-OCT-2000.
XX
PF 01-APR-1999; 99WO-EP002283.
XX
PR 01-APR-1999; 99WO-EP002283.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 2000-665002/64.
XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding fragments.
XX
PS Disclosure; Page 15; 68pp; English.
XX
CC The present invention is concerned with producing scaffold proteins based
CC upon the human CTLA-4 SCA domain. These scaffold proteins can be used as
CC a scaffold to bind antigen- or receptor-binding fragments. These can be
CC used in the treatment of diseases such as cancer, atherosclerosis,
CC thrombosis, osteoporosis, rheumatoid arthritis and diabetic retinopathy.
CC Sequences AAB29930-B29939 were used in the production of the proteins of
CC the invention
XX
SQ Sequence 7 AA;

AAB30075
ID AAB30075 standard; peptide; 7 AA.
XX
AC AAB30075;
XX
DT 09-FEB-2001 (first entry)
XX
DE Scaffold protein SCA S4 peptide SEQ ID NO: 136.
XX
KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.
XX
OS Synthetic.
XX
PN WO200060070-A1.
XX
PD 12-OCT-2000.
XX
PF 01-APR-1999; 99WO-EP002283.
XX
PR 01-APR-1999; 99WO-EP002283.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 2000-665002/64.
XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding fragments.
XX
PS Disclosure; Page 15; 68pp; English.
XX
CC The present invention is concerned with producing scaffold proteins based
CC upon the human CTLA-4 SCA domain. These scaffold proteins can be used as
CC a scaffold to bind antigen- or receptor-binding fragments. These can be
CC used in the treatment of diseases such as cancer, atherosclerosis,
CC thrombosis, osteoporosis, rheumatoid arthritis and diabetic retinopathy.
CC Sequences AAB29930-B29939 were used in the production of the proteins of
CC the invention
XX
SQ Sequence 7 AA;

Query Match 78.1%; Score 25; DB 3; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLTIS 6
:|||||
Db 2 YTLTIS 7

RESULT 5
AAY40736
ID AAY40736 standard; peptide; 7 AA.
XX
AC AAY40736;
XX
DT 01-DEC-1999 (first entry)
XX
DE S4 derivative #10, beta strand of scaffold protein structure.
XX
KW Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.
XX
OS Synthetic.
XX
PN EP947582-A1.
XX
PD 06-OCT-1999.
XX
PF 31-MAR-1998; 98EP-00870065.
XX
PR 31-MAR-1998; 98EP-00870065.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 1999-542958/46.
XX
PT New scaffold protein, useful for stabilizing antigens used as vaccines.
XX
PS Disclosure; Page 6; 105pp; English.
XX
CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides (AAY40601-
CC Y40609) together form a single-chain scaffold protein which contains at
CC least 1 disulfide bond, contains less than 10% alpha helix and contains
CC at least 6 beta-strands. The scaffold protein is constructed of beta
CC strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines
XX
SQ Sequence 7 AA;

Query Match 75.0%; Score 24; DB 2; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLTIS 6
|||:||
Db 2 FTLSIS 7

RESULT 6
AAB30074
ID AAB30074 standard; peptide; 7 AA.
XX
AC AAB30074;
XX
DT 09-FEB-2001 (first entry)
XX
DE Scaffold protein SCA S4 peptide SEQ ID NO: 135.
XX
KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.
XX
OS Synthetic.
XX
PN WO200060070-A1.
XX
PD 12-OCT-2000.
XX
PF 01-APR-1999; 99WO-EP002283.
XX
PR 01-APR-1999; 99WO-EP002283.

CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines
XX
SQ Sequence 7 AA;

Query Match 87.5%; Score 28; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLTIS 6
||||||
Db 2 FTLTIS 7

RESULT 2
AAB30076
ID AAB30076 standard; peptide; 7 AA.
XX
AC AAB30076;
XX
DT 09-FEB-2001 (first entry)
XX
DE Scaffold protein SCA S4 peptide SEQ ID NO: 137.
XX
KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.
XX
OS Synthetic.
XX
PN WO200060070-A1.
XX
PD 12-OCT-2000.
XX
PF 01-APR-1999; 99WO-EP002283.
XX
PR 01-APR-1999; 99WO-EP002283.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 2000-665002/64.
XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding fragments.
XX
PS Disclosure; Page 15; 68pp; English.
XX
CC The present invention is concerned with producing scaffold proteins based
CC upon the human CTLA-4 SCA domain. These scaffold proteins can be used as
CC a scaffold to bind antigen- or receptor-binding fragments. These can be
CC used in the treatment of diseases such as cancer, atherosclerosis,
CC thrombosis, osteoporosis, rheumatoid arthritis and diabetic retinopathy.
CC Sequences AAB29930-B29939 were used in the production of the proteins of
CC the invention
XX
SQ Sequence 7 AA;

Query Match 87.5%; Score 28; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLTIS 6
||||||
Db 2 FTLTIS 7

RESULT 3
AAY40737
ID AAY40737 standard; peptide; 7 AA.
XX
AC AAY40737;
XX
DT 01-DEC-1999 (first entry)
XX
DE S4 derivative #11, beta strand of scaffold protein structure.
XX
KW Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.
XX
OS Synthetic.
XX
PN EP947582-A1.
XX
PD 06-OCT-1999.
XX
PF 31-MAR-1998; 98EP-00870065.
XX
PR 31-MAR-1998; 98EP-00870065.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 1999-542958/46.
XX
PT New scaffold protein, useful for stabilizing antigens used as vaccines.
XX
PS Disclosure; Page 6; 105pp; English.
XX
CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides (AAY40601-
CC Y40609) together form a single-chain scaffold protein which contains at
CC least 1 disulfide bond, contains less than 10% alpha helix and contains
CC at least 6 beta-strands. The scaffold protein is constructed of beta
CC strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines
XX
SQ Sequence 7 AA;

Query Match 78.1%; Score 25; DB 2; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLTIS 6
:|||||
Db 2 YTLTIS 7

RESULT 4

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:40:31 ; Search time 45.6667 Seconds
(without alignments)
43.310 Million cell updates/sec

Title: US-09-712-819C-1
Perfect score: 32
Sequence: 1 FTLTISS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 92273

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	87.5	7	2	AAY40738	Aay40738 S4 deriva
2	28	87.5	7	3	AAB30076	Aab30076 Scaffold
3	25	78.1	7	2	AAY40737	Aay40737 S4 deriva
4	25	78.1	7	3	AAB30075	Aab30075 Scaffold
5	24	75.0	7	2	AAY40736	Aay40736 S4 deriva
6	24	75.0	7	3	AAB30074	Aab30074 Scaffold
7	21	65.6	7	2	AAY40735	Aay40735 S4 deriva
8	21	65.6	7	3	AAB30073	Aab30073 Scaffold
9	20	62.5	5	5	AAU85454	Aau85454 Human col
10	19	59.4	6	4	AAM51422	Aam51422 Integrin
11	18	56.2	5	6	ABR55418	Abr55418 Amino aci
12	18	56.2	7	2	AAW58711	Aaw58711 Tryptic 4
13	18	56.2	7	3	AAY64225	Aay64225 Cadherin-
14	18	56.2	7	5	ABB84102	Abb84102 Human sin
15	17	53.1	5	6	ABU12234	Abu12234 Streptoco
16	17	53.1	5	6	ABU12235	Abu12235 Streptoco
17	17	53.1	6	2	AAW39453	Aaw39453 Human T c
18	17	53.1	6	2	AAW75358	Aaw75358 Hexapepti
19	17	53.1	6	2	AAW75290	Aaw75290 Hexapepti
20	17	53.1	6	6	AAE31832	Aae31832 Androgen
21	17	53.1	7	2	AAR07656	Aar07656 Ribonucle
22	17	53.1	7	2	AAY40723	Aay40723 S3 deriva
23	17	53.1	7	3	AAB30062	Aab30062 Scaffold
24	17	53.1	7	4	AAM44200	Aam44200 H11 bindi
25	17	53.1	7	4	AAM44205	Aam44205 H11 bindi
26	16	50.0	6	1	AAP93345	Aap93345 Portion o
27	16	50.0	6	2	AAR93089	Aar93089 GAL4 DNA
28	16	50.0	6	2	AAW61443	Aaw61443 Gal 4 pro
29	16	50.0	6	2	AAW76753	Aaw76753 Murine si
30	16	50.0	6	2	AAW92487	Aaw92487 Murine Ga
31	16	50.0	6	3	AAY90491	Aay90491 GAL4 DNA
32	16	50.0	6	3	AAB03571	Aab03571 Nuclear c
33	16	50.0	6	4	AAB36961	Aab36961 Peptide #
34	16	50.0	6	4	AAB60836	Aab60836 Peptide #
35	16	50.0	6	5	ABJ11455	Abj11455 Human 125
36	16	50.0	6	5	AAE28111	Aae28111 Human imm
37	16	50.0	6	6	ABU72693	Abu72693 Novel pro
38	16	50.0	6	6	ABR45358	Abr45358 Staphyloc
39	16	50.0	6	6	ABR46478	Abr46478 Staphyloc
40	16	50.0	6	6	ABR44966	Abr44966 Staphyloc
41	16	50.0	6	6	ABR45694	Abr45694 Staphyloc
42	16	50.0	6	6	ABR46870	Abr46870 Staphyloc
43	16	50.0	6	6	ABR46086	Abr46086 Staphyloc
44	16	50.0	6	6	ABJ57003	Abj57003 184P1E2-r
45	16	50.0	7	2	AAR72775	Aar72775 Mammalian

ALIGNMENTS

RESULT 1
AAY40738
ID AAY40738 standard; peptide; 7 AA.
XX
AC AAY40738;
XX
DT 01-DEC-1999 (first entry)
XX
DE S4 derivative #12, beta strand of scaffold protein structure.
XX
KW Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.
XX
OS Synthetic.
XX
PN EP947582-A1.
XX
PD 06-OCT-1999.
XX
PF 31-MAR-1998; 98EP-00870065.
XX
PR 31-MAR-1998; 98EP-00870065.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 1999-542958/46.
XX
PT New scaffold protein, useful for stabilizing antigens used as vaccines.
XX
PS Disclosure; Page 6; 105pp; English.
XX
CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides (AAY40601-
CC Y40609) together form a single-chain scaffold protein which contains at
CC least 1 disulfide bond, contains less than 10% alpha helix and contains
CC at least 6 beta-strands. The scaffold protein is constructed of beta
CC strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens

```
Qy          3 LT 4
              :|
Db          1 IT 2


RESULT 14
P83492
ID   P83492      PRELIMINARY;      PRT;     7 AA.
AC   P83492;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Alkaline protease Gr3 (EC 3.4.21.-) (Fragment).
OS   Bionectria ochroleuca (Gliocladium roseum).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Bionectriaceae; Bionectria.
OX   NCBI_TaxID=29856;
RN   [1]
RP   SEQUENCE, FUNCTION, AND SUBCELLULAR LOCATION.
RC   STRAIN=Gr87;
RA   Zhao M., Zhang K.;
RL   Submitted (DEC-2002) to Swiss-Prot.
CC   -!- FUNCTION: ACTS AS A SERINE PROTEASE.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
DR   InterPro; IPR000209; Peptidase_S8.
DR   PROSITE; PS00136; SUBTILASE_ASP; PARTIAL.
DR   PROSITE; PS00137; SUBTILASE_HIS; PARTIAL.
DR   PROSITE; PS00138; SUBTILASE_SER; PARTIAL.
KW   Hydrolase; Serine protease.
FT   NON_TER       7       7
SQ   SEQUENCE   7 AA;  688 MW;  776DD455A6C1ADB0 CRC64;

  Query Match             21.9%;  Score 7;  DB 3;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 TISS 7
              | |:
Db          2 TQSN 5


RESULT 15
Q65578
ID   Q65578      PRELIMINARY;      PRT;     7 AA.
AC   Q65578;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Bovine herpesvirus 1.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=10320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Cooper;
RX   MEDLINE=95313343; PubMed=7793062;
RA   Vlcek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
RA   Letchworth G.J., Schwyzer M.;
RT   "Nucleotide sequence analysis of a 30-kb region of the bovine
RT   herpesvirus 1 genome which exhibits a colinear gene arrangement with
RT   the UL21 to UL4 genes of herpes simplex virus.";
RL   Virology 210:100-108(1995).
DR   EMBL; Z48053; CAA88130.1; -.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   7 AA;  758 MW;  6DD33455B1F1B1C0 CRC64;

  Query Match             21.9%;  Score 7;  DB 12;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 TISS 7
              | |
Db          1 TTHS 4


Search completed: June  1, 2004, 11:18:36
Job time : 33.3333 secs
```

```
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 50S ribosomal protein L10 beta (Fragment).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435798; PubMed=10874046;
RA   Yamaguchi K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the 50 S subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 275:28466-28482(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC   -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002363; Ribosomal_L10eub.
DR   PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  675 MW;  6321B415B05DB000 CRC64;

  Query Match             25.0%;  Score 8;  DB 10;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 IS 6
              ||
Db          2 IS 3


RESULT 11
P82182
ID   P82182      PRELIMINARY;      PRT;     6 AA.
AC   P82182;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 50S ribosomal protein L10 gamma (Fragment).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435798; PubMed=10874046;
RA   Yamaguchi K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the 50 S subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 275:28466-28482(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC   -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002363; Ribosomal_L10eub.
DR   PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  675 MW;  6321B415B05DB000 CRC64;

  Query Match             25.0%;  Score 8;  DB 10;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 IS 6
              ||
Db          2 IS 3


RESULT 12
P72081
ID   P72081      PRELIMINARY;      PRT;     7 AA.
AC   P72081;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   3'-methylcephem hydroxylase (Fragment).
GN   CEFF.
OS   Nocardia lactamdurans.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Pseudonocardineae; Pseudonocardiaceae; Amycolatopsis.
OX   NCBI_TaxID=1913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96009872; PubMed=7557411;
RA   Coque J., Perez-Llarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,
RA   Liras P.;
RT   "Characterization of the cmcH genes of Nocardia lactamdurans and
RT   Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
RT   O-carbamoyltransferase for cephamycin biosynthesis.";
RL   Gene 162:21-27(1995).
DR   EMBL; Z21682; CAA79797.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  746 MW;  75A1B2CDD1E681C0 CRC64;

  Query Match             25.0%;  Score 8;  DB 2;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 1e+06;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ISS 7
              ::|
Db          5 VTS 7


RESULT 13
Q47029
ID   Q47029      PRELIMINARY;      PRT;     7 AA.
AC   Q47029;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Aad A1 protein (Fragment).
GN   AAD A1.
OS   Enterobacter cloacae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Enterobacter.
OX   NCBI_TaxID=550;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94079349; PubMed=8257126;
RA   Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;
RT   "Analysis of the aac(3)-VIa gene encoding a novel 3-N-
RT   acetyltransferase.";
RL   Antimicrob. Agents Chemother. 37:2074-2079(1993).
DR   EMBL; M88012; AAA16193.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  744 MW;  633862D2C321A030 CRC64;

  Query Match             21.9%;  Score 7;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
RESULT 6
P83533
ID   P83533      PRELIMINARY;      PRT;     6 AA.
AC   P83533;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Unknown protein from 2D-page (Fragment).
OS   Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1625;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=DSM 20451;
RX   PubMed=12112860;
RA   Drews O., Weiss W., Reil G., Parlar H., Wait R., Goerg A.;
RT   "High pressure effects step-wise altered protein expression in
RT   Lactobacillus sanfranciscensis.";
RL   Proteomics 2:765-774(2002).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS UNKNOWN
CC       PROTEIN IS: 15 KDA.
FT   NON_TER       1      1
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  590 MW;  6DDDD452D1AAC000 CRC64;

  Query Match             31.2%;  Score 10;  DB 2;  Length 6;
  Best Local Similarity   25.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TISS 7
              |:::
Db          2 TVNA 5


RESULT 7
O07354
ID   O07354      PRELIMINARY;      PRT;     7 AA.
AC   O07354;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NifK (Fragment).
GN   NIFK.
OS   Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothece PCC 8801).
OC   Bacteria; Cyanobacteria; Chroococcales; Cyanothece.
OX   NCBI_TaxID=41431;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RF-1;
RX   MEDLINE=99231861; PubMed=10217509;
RA   Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RT   "Organization and expression of nitrogen-fixation genes in the aerobic
RT   nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
RT   RF-1.";
RL   Microbiology 145:743-753(1999).
DR   EMBL; AF003700; AAC35193.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  849 MW;  7412C72AA9D5B030 CRC64;

  Query Match             28.1%;  Score 9;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          3 FDL 5


RESULT 8
P82445
ID   P82445      PRELIMINARY;      PRT;     7 AA.
AC   P82445;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   10 kDa cell wall protein (Fragment).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. PETIT HAVANA;
RA   Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
RA   Wojtaszek P., Bolwell G.P.;
RT   "Proteomic study of secondary cell wall proteins from transformed
RT   tobacco culture.";
RL   Planta 0:0-0(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL.
CC   -!- TISSUE SPECIFICITY: XYLEM.
DR   GO; GO:0005618; C:cell wall; IEA.
KW   Cell wall.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  758 MW;  69D2C1E862D1B2A0 CRC64;

  Query Match             28.1%;  Score 9;  DB 10;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 1e+06;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTI 5
              :|:
Db          1 VTV 3


RESULT 9
Q8JE81
ID   Q8JE81      PRELIMINARY;      PRT;     7 AA.
AC   Q8JE81;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Truncated pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=4874;
RX   MEDLINE=22056123; PubMed=12060770;
RA   Beerenwinkel N., Schmidt B., Walter H., Kaiser R., Lengauer T.,
RA   Hoffmann D., Korn K., Selbig J.;
RT   "Diversity and complexity of HIV-1 drug resistance: A bioinformatics
RT   approach to predicting phenotype from genotype.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:8271-8276(2002).
DR   EMBL; AF347267; AAK32344.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  885 MW;  76C37731A046C700 CRC64;

  Query Match             28.1%;  Score 9;  DB 15;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TL 3
              ||
Db          4 TL 5


RESULT 10
P82181
ID   P82181      PRELIMINARY;      PRT;     6 AA.
AC   P82181;
```

```
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Catechol-2,3-dioxygenase (Fragment).
GN   PHNE.
OS   Sphingomonas chungbukensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC   Sphingomonadaceae; Sphingomonas.
OX   NCBI_TaxID=56193;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DJ77;
RA   Kim Y.-C.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U88298; AAB66311.1; -.
DR   GO; GO:0016702; F:oxidoreductase activity, acting on single d. . .; IEA.
KW   Dioxygenase.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  868 MW;  71A452D1A699D460 CRC64;

  Query Match             37.5%;  Score 12;  DB 2;  Length 7;
  Best Local Similarity   20.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTISS 7
              :|:::
Db          3 MTVNT 7


RESULT 3
Q7Z1C0
ID   Q7Z1C0      PRELIMINARY;      PRT;     7 AA.
AC   Q7Z1C0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein W01B11.6.
GN   W01B11.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Wilson R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Bradshaw H., Graves T., Blair T.;
RT   "The sequence of C. elegans cosmid W01B11.";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Wilson R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF043704; AAO38592.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   7 AA;  874 MW;  72D1A9DB5041A6F0 CRC64;

  Query Match             37.5%;  Score 12;  DB 5;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              ||:
Db          5 FTV 7


RESULT 4
Q54248
ID   Q54248      PRELIMINARY;      PRT;     7 AA.
AC   Q54248;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RplO protein (Fragment).
GN   RPLO.
OS   Streptomyces griseus.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1911;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2-3-11;
RX   MEDLINE=20011291; PubMed=10542330;
RA   Poehling S., Piepersberg W., Wehmeier U.F.;
RT   "Analysis and regulation of the sec Y gene from Streptomyces griseus
RT   N2-3-11 and interaction of the SecY protein with the SecA protein.";
RL   Biochim. Biophys. Acta 1447:298-302(1999).
DR   EMBL; X95915; CAA65160.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  760 MW;  72C72B01B2D1B2A0 CRC64;

  Query Match             34.4%;  Score 11;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLT 4
              |:|
Db          2 TVT 4


RESULT 5
P83530
ID   P83530      PRELIMINARY;      PRT;     7 AA.
AC   P83530;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Unknown protein from 2D-page (Fragment).
OS   Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1625;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=DSM 20451;
RX   PubMed=12112860;
RA   Drews O., Weiss W., Reil G., Parlar H., Wait R., Goerg A.;
RT   "High pressure effects step-wise altered protein expression in
RT   Lactobacillus sanfranciscensis.";
RL   Proteomics 2:765-774(2002).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS UNKNOWN
CC       PROTEIN IS: 15 KDA.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  790 MW;  6AAB02CAA731B2A0 CRC64;

  Query Match             34.4%;  Score 11;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLTI 5
              || :
Db          2 TLDV 5
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:53:46 ; Search time 31.3333 Seconds
(without alignments)
70.488 Million cell updates/sec

Title: US-09-712-819C-1
Perfect score: 32
Sequence: 1 FTLTISS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 74

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	16	50.0	7	2	P70804	P70804 azotobacter
2	12	37.5	7	2	O34028	O34028 sphingomona
3	12	37.5	7	5	Q7Z1C0	Q7z1c0 caenorhabdi
4	11	34.4	7	2	Q54248	Q54248 streptomyce
5	11	34.4	7	2	P83530	P83530 lactobacill
6	10	31.2	6	2	P83533	P83533 lactobacill
7	9	28.1	7	2	O07354	O07354 synechococc
8	9	28.1	7	10	P82445	P82445 nicotiana t
9	9	28.1	7	15	Q8JE81	Q8je81 human immun
10	8	25.0	6	10	P82181	P82181 spinacia ol
11	8	25.0	6	10	P82182	P82182 spinacia ol
12	8	25.0	7	2	P72081	P72081 nocardia la
13	7	21.9	7	2	Q47029	Q47029 enterobacte
14	7	21.9	7	3	P83492	P83492 bionectria
15	7	21.9	7	12	Q65578	Q65578 bovine herp
16	6	18.8	5	13	P83308	P83308 gallus gall
17	6	18.8	6	10	P82541	P82541 spinacia ol
18	6	18.8	7	2	O50556	O50556 actinobacil
19	6	18.8	7	2	Q8KMS9	Q8kms9 enterobacte
20	6	18.8	7	4	Q15903	Q15903 homo sapien
21	6	18.8	7	8	O98866	O98866 spinacia ol
22	6	18.8	7	8	Q95945	Q95945 saccharomyc
23	6	18.8	7	10	Q9C5B3	Q9c5b3 arabidopsis
24	6	18.8	7	11	Q8K3H6	Q8k3h6 rattus norv
25	6	18.8	7	12	Q9YQ10	Q9yq10 transmissib
26	6	18.8	7	13	O42564	O42564 fugu rubrip
27	5	15.6	5	2	P83073	P83073 bacillus ce
28	5	15.6	7	2	Q8KMS3	Q8kms3 klebsiella
29	5	15.6	7	2	Q47505	Q47505 escherichia
30	5	15.6	7	2	Q8GL12	Q8gl12 borrelia bu
31	5	15.6	7	2	Q8GL04	Q8gl04 borrelia bu
32	5	15.6	7	8	Q8MFY6	Q8mfy6 taraxacum (
33	5	15.6	7	10	P93233	P93233 lycopersico
34	5	15.6	7	11	Q63668	Q63668 rattus norv
35	5	15.6	7	12	Q67113	Q67113 influenzavi
36	5	15.6	7	15	Q07624	Q07624 rous sarcom
37	4	12.5	4	5	P83568	P83568 sepia offic
38	4	12.5	4	11	Q08433	Q08433 rattus sp.
39	4	12.5	6	5	P83569	P83569 sepia offic
40	4	12.5	7	2	Q8GL00	Q8gl00 borrelia bu
41	4	12.5	7	4	Q8NHH7	Q8nhh7 homo sapien
42	4	12.5	7	4	Q15897	Q15897 homo sapien
43	4	12.5	7	8	P92214	P92214 amblyopyrum
44	4	12.5	7	8	P92393	P92393 hordeum vul
45	4	12.5	7	8	P92403	P92403 lophopyrum

ALIGNMENTS

RESULT 1
P70804
ID P70804 PRELIMINARY; PRT; 7 AA.
AC P70804;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE AlgT protein (Fragment).
GN ALGT.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E;
RX MEDLINE=96427318; PubMed=8830682;
RA Rehm B.H.A., Ertesvag H., Valla S.;
RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is
RT part of an alg gene cluster physically organized in a manner similar
RT to that in Pseudomonas aeruginosa.";
RL J. Bacteriol. 178:5884-5889(1996).
DR EMBL; X87973; CAA61230.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 50.0%; Score 16; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 1e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TISS 7
|:||
Db 2 TVSS 5

RESULT 2
O34028
ID O34028 PRELIMINARY; PRT; 7 AA.
AC O34028;

Query Match 25.0%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTL 3
              | |
Db          5 FGL 7
```

RESULT 14
ALL5_CARMA
ID ALL5_CARMA STANDARD; PRT; 7 AA.
AC P81808;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 5.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC -!- SIMILARITY: Belongs to the allatostatin family.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 781 MW; 672879CDCB476420 CRC64;

Query Match 25.0%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTL 3
              | |
Db          5 FGL 7
```

RESULT 15
ALL7_CYDPO
ID ALL7_CYDPO STANDARD; PRT; 7 AA.
AC P82158;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 7.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricoidea; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -!- SIMILARITY: Belongs to the allatostatin family.
KW Neuropeptide; Amidation.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 873 MW; 672879CABB569350 CRC64;

Query Match 25.0%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTL 3
              | |
Db          5 FGL 7
```

Search completed: June 1, 2004, 11:16:49
Job time : 9.66667 secs

```
Qy          1 FT 2
              :|
Db          3 YT 4


RESULT 10
LOK1_LOCMI
ID   LOK1_LOCMI     STANDARD;      PRT;     6 AA.
AC   P41491;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Locustakinin I.
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC   Acridoidea; Acrididae; Oedipodinae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=92262851; PubMed=1585017;
RA   Schoofs L., Holman G.M., Proost P., van Damme J., Hayes T.K.,
RA   de Loof A.;
RT   "Locustakinin, a novel myotropic peptide from Locusta migratoria,
RT   isolation, primary structure and synthesis.";
RL   Regul. Pept. 37:49-57(1992).
CC   -!- FUNCTION: Myotropic peptide. May be important in the stimulation
CC       of ion transport and inhibition of diuretic activity in Malpighian
CC       tubules.
CC   -!- SUBCELLULAR LOCATION: Secreted.
DR   PIR; A61068; A61068.
KW   Neuropeptide; Amidation.
FT   MOD_RES       6      6       AMIDATION.
SQ   SEQUENCE   6 AA;  654 MW;  686365A5B9CDB000 CRC64;

  Query Match             25.0%;  Score 8;  DB 1;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 SS 7
              ||
Db          3 SS 4


RESULT 11
ALL2_CARMA
ID   ALL2_CARMA     STANDARD;      PRT;     7 AA.
AC   P81805;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 2.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       7      7       AMIDATION (POTENTIAL).
SQ   SEQUENCE   7 AA;  770 MW;  672879CDCB5DDB70 CRC64;

  Query Match             25.0%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          5 FGL 7


RESULT 12
ALL3_CARMA
ID   ALL3_CARMA     STANDARD;      PRT;     7 AA.
AC   P81806;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 3.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Multigene family.
SQ   SEQUENCE   7 AA;  796 MW;  672879CDCB476B70 CRC64;

  Query Match             25.0%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          5 FGL 7


RESULT 13
ALL4_CARMA
ID   ALL4_CARMA     STANDARD;      PRT;     7 AA.
AC   P81807;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 4.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Multigene family.
SQ   SEQUENCE   7 AA;  782 MW;  672879CDCB476AC0 CRC64;
```

```
RT   "Isolation and structure of the Streptococcus faecalis sex pheromone,
RT   cAM373.";
RL   FEBS Lett. 206:69-72(1986).
CC   -!- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
CC       HARBORING PAM373.
CC   -!- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
CC       SPECIFICITY OF PHEROMONES TO PLASMIDS.
CC   -!- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
DR   PIR; A25269; A25269.
KW   Pheromone.
SQ   SEQUENCE   7 AA;  734 MW;  75BDD72059C05DB0 CRC64;

  Query Match             28.1%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          3 FIL 5


RESULT 7
AL14_CARMA
ID   AL14_CARMA     STANDARD;      PRT;     5 AA.
AC   P81817;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 14.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       5       5       AMIDATION (POTENTIAL).
SQ   SEQUENCE   5 AA;  586 MW;  672879D5AB300000 CRC64;

  Query Match             25.0%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          3 FGL 5


RESULT 8
BIOB_CITFR
ID   BIOB_CITFR     STANDARD;      PRT;     5 AA.
AC   P12997;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Biotin synthase (EC 2.8.1.6) (Biotin synthetase) (Fragment).
GN   BIOB.
OS   Citrobacter freundii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Citrobacter.
OX   NCBI_TaxID=546;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89006280; PubMed=2971595;
RA   Shiuan D., Campbell A.;
RT   "Transcriptional regulation and gene arrangement of Escherichia coli,
RT   Citrobacter freundii and Salmonella typhimurium biotin operons.";
RL   Gene 67:203-211(1988).
CC   -!- CATALYTIC ACTIVITY: Dethiobiotin + sulfur = biotin.
CC   -!- PATHWAY: Biotin biosynthesis; last step.
CC   -!- SIMILARITY: Belongs to the biotin and lipoic acid synthetases
CC       family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M21922; -; NOT_ANNOTATED_CDS.
DR   PIR; I40698; I40698.
KW   Biotin biosynthesis; Iron-sulfur; Transferase.
FT   NON_TER       5       5
SQ   SEQUENCE   5 AA;  532 MW;  75A5B1EDD6F00000 CRC64;

  Query Match             25.0%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 SS 7
              ||
Db          4 SS 5


RESULT 9
PSK_DAUCA
ID   PSK_DAUCA      STANDARD;      PRT;     5 AA.
AC   P58261;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Phytosulfokine-alpha (PSK-alpha) [Contains: Phytosulfokine-beta (PSK-
DE   beta)].
OS   Daucus carota (Carrot).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   campanulids; Apiales; Apiaceae; Apioideae; Scandiceae; Daucinae;
OC   Daucus.
OX   NCBI_TaxID=4039;
RN   [1]
RP   SEQUENCE, AND IDENTIFICATION BY MASS SPECTROMETRY.
RC   STRAIN=cv. US-Harumakigosun;
RX   MEDLINE=20212743; PubMed=10750705;
RA   Hanai H., Matsuno T., Yamamoto M., Matsubayashi Y., Kobayashi T.,
RA   Kamada H., Sakagami Y.;
RT   "A secreted peptide growth factor, phytosulfokine, acting as a
RT   stimulatory factor of carrot somatic embryo formation.";
RL   Plant Cell Physiol. 41:27-32(2000).
CC   -!- FUNCTION: In presence of 2,4-D, stimulates proliferation of the
CC       cells, but does not stimulate differentiation into the somatic
CC       embryos.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: Sulfation is important for activity and for the binding to a
CC       putative membrane receptor (By similarity).
CC   -!- SIMILARITY: Belongs to the phytosulfokine family.
KW   Growth factor; Sulfation.
FT   PEPTIDE       1       4       PHYTOSULFOKINE-BETA.
FT   MOD_RES       1       1       SULFATION.
FT   MOD_RES       3       3       SULFATION.
SQ   SEQUENCE   5 AA;  687 MW;  76C1BB504B300000 CRC64;

  Query Match             25.0%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
RT   "The structure of new peptides from the Australin red tree frog
RT   'Litoria rubella'. The skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC   -!- MASS SPECTROMETRY: MW=655; METHOD=FAB.
KW   Amphibian defense peptide; Amidation.
FT   MOD_RES       5       5       AMIDATION.
SQ   SEQUENCE   5 AA;  656 MW;  71A9C9CB10300000 CRC64;

  Query Match             34.4%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FT 2
               ||
Db          4 FT 5


RESULT 3
RE32_LITRU
ID   RE32_LITRU     STANDARD;      PRT;     5 AA.
AC   P82073;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Rubellidin 3.2.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Skin secretion;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litori electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:639-645(1999).
CC   -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
KW   Amphibian defense peptide.
SQ   SEQUENCE   5 AA;  570 MW;  71A9C9C862A00000 CRC64;

  Query Match             34.4%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FT 2
               ||
Db          4 FT 5


RESULT 4
EI04_LITRU
ID   EI04_LITRU     STANDARD;      PRT;     5 AA.
AC   P82100;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Electrin 4.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Skin secretion;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litori electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:639-645(1999).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Skin.
KW   Amphibian defense peptide; Amidation.
FT   MOD_RES       5       5       AMIDATION.
SQ   SEQUENCE   5 AA;  616 MW;  61F2D1A059A00000 CRC64;

  Query Match             31.2%;  Score 10;  DB 1;  Length 5;
  Best Local Similarity   33.3%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTI 5
               :|:
Db          2 ITV 4


RESULT 5
CCF1_ENTFA
ID   CCF1_ENTFA     STANDARD;      PRT;     7 AA.
AC   P20104;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cCF10.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=89008313; PubMed=3139658;
RA   Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,
RA   Adsit J.C., Dunny G.M., Suzuki A.;
RT   "Structure of cCF10, a peptide sex pheromone which induces
RT   conjugative transfer of the Streptococcus faecalis tetracycline
RT   resistance plasmid, pCF10.";
RL   J. Biol. Chem. 263:14574-14578(1988).
CC   -!- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       HEMOLYSIN PLASMID PCF10.
DR   PIR; A30812; A30812.
KW   Pheromone.
SQ   SEQUENCE   7 AA;  790 MW;  72C9D2C731B2C740 CRC64;

  Query Match             28.1%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TL 3
               ||
Db          3 TL 4


RESULT 6
CIA_ENTFA
ID   CIA_ENTFA      STANDARD;      PRT;     7 AA.
AC   P11932;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cAM373 (Clumping-inducing agent) (CIA).
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=87005252; PubMed=3093276;
RA   Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
RA   White B.A., An F.Y., Clewell D.B., Suzuki A.;
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:41:50 ; Search time 7.66667 Seconds
(without alignments)
47.542 Million cell updates/sec

Title: US-09-712-819C-1
Perfect score: 32
Sequence: 1 FTLTISS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 88

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	12	37.5	7	1	GFRP_MOUSE	P99025 mus musculu
2	11	34.4	5	1	RE31_LITRU	P82072 litoria rub
3	11	34.4	5	1	RE32_LITRU	P82073 litoria rub
4	10	31.2	5	1	EI04_LITRU	P82100 litoria rub
5	9	28.1	7	1	CCF1_ENTFA	P20104 enterococcu
6	9	28.1	7	1	CIA_ENTFA	P11932 enterococcu
7	8	25.0	5	1	AL14_CARMA	P81817 carcinus ma
8	8	25.0	5	1	BIOB_CITFR	P12997 citrobacter
9	8	25.0	5	1	PSK_DAUCA	P58261 daucus caro
10	8	25.0	6	1	LOK1_LOCMI	P41491 locusta mig
11	8	25.0	7	1	ALL2_CARMA	P81805 carcinus ma
12	8	25.0	7	1	ALL3_CARMA	P81806 carcinus ma
13	8	25.0	7	1	ALL4_CARMA	P81807 carcinus ma
14	8	25.0	7	1	ALL5_CARMA	P81808 carcinus ma
15	8	25.0	7	1	ALL7_CYDPO	P82158 cydia pomon
16	8	25.0	7	1	UC24_MAIZE	P80630 zea mays (m
17	7	21.9	4	1	EOSI_HUMAN	P02731 homo sapien
18	7	21.9	4	1	RM01_YEAST	P36515 saccharomyc
19	7	21.9	5	1	BIOA_CITFR	P13071 citrobacter
20	7	21.9	6	1	UN06_CLOPA	P81351 clostridium
21	6	18.8	4	1	ACH1_ACHFU	P35904 achatina fu
22	6	18.8	4	1	FAR3_HIRME	P42562 hirudo medi
23	6	18.8	4	1	FAR4_HIRME	P42563 hirudo medi
24	6	18.8	4	1	FFKA_ANTEL	P58705 anthopleura
25	6	18.8	4	1	FLRF_HIRME	P42561 hirudo medi
26	6	18.8	4	1	FLRN_ANTEL	P58707 anthopleura
27	6	18.8	4	1	FMRF_MACNI	P01162 macrocallis
28	6	18.8	4	1	FYRI_ANTEL	P58706 anthopleura
29	6	18.8	4	1	OCP1_OCTMI	P58648 octopus min
30	6	18.8	5	1	EI03_LITRU	P82099 litoria rub
31	6	18.8	5	1	FARP_ARTTR	P41853 artioposthi
32	6	18.8	5	1	PAP2_PARMA	P81864 pardachirus
33	6	18.8	5	1	RE11_LITRU	P82070 litoria rub
34	6	18.8	5	1	RE21_LITRU	P82071 litoria rub
35	6	18.8	5	1	SUGA_ACHDO	P19991 acheta dome
36	6	18.8	5	1	TPIS_CANFA	P54714 canis famil
37	6	18.8	5	1	UC22_MAIZE	P80628 zea mays (m
38	6	18.8	6	1	CIP1_MYTED	P13736 mytilus edu
39	6	18.8	6	1	CIP2_MYTED	P13737 mytilus edu
40	6	18.8	6	1	EI01_LITRU	P82096 litoria rub
41	6	18.8	6	1	FARP_MONEX	P41966 moniezia ex
42	6	18.8	7	1	FAR1_ASCSU	P31889 ascaris suu
43	6	18.8	7	1	FAR1_HELTI	P41871 helisoma tr
44	6	18.8	7	1	FAR1_MACRS	P83274 macrobrachi
45	6	18.8	7	1	FAR1_PROCL	P38499 procambarus

ALIGNMENTS

RESULT 1
GFRP_MOUSE
ID GFRP_MOUSE STANDARD; PRT; 7 AA.
AC P99025;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
GN GCHFR OR GFRP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA Cowthorne M.;
RL Submitted (AUG-1998) to Swiss-Prot.
CC -!- FUNCTION: Mediates tetrahydrobiopterin inhibition of GTP
CC cyclohydrolase I. This inhibition is reversed by L-phenylalanine
CC (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
DR SWISS-2DPAGE; P99025; MOUSE.
FT INIT_MET 0 0
FT NON_TER 7 7
SQ SEQUENCE 7 AA; 806 MW; 71B5B057273B4700 CRC64;

Query Match 37.5%; Score 12; DB 1; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          3 LTISS 7
              | ||:
Db          3 LLIST 7
```

RESULT 2
RE31_LITRU
ID RE31_LITRU STANDARD; PRT; 5 AA.
AC P82072;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rubellidin 3.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC Pelodryadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Skin secretion;
RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;

```
Qy          5 ISS 7
              ||:
Db          1 IST 3


RESULT 15
A43766
28K ubiquitin-immunoreactive protein - inky cap (Coprinus cinereus) (fragment)
C;Species: Coprinus cinereus
C;Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 30-Sep-1993
C;Accession: A43766
R;Kanda, T.; Inoue, M.; Akiyama, M.
Biochimie 72, 355-359, 1990
A;Title: Purification and characterization of an ubiquitin-immuno-reactive protein local
A;Reference number: A43766; MUID:91002724; PMID:1698461
A;Accession: A43766
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-6 <KAN>

  Query Match             28.1%;  Score 9;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TL 3
              ||
Db          5 TL 6


Search completed: June  1, 2004, 11:19:23
Job time : 12.6667 secs
```

```
              : :|
Db          2 YPIT 5
```

RESULT 9
A37114
hypoxanthine phosphoribosyltransferase (EC 2.4.2.8) - fluke (Schistosoma mansoni) (fragm
C;Species: Schistosoma mansoni
C;Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 23-Jun-1993
C;Accession: A37114
R;Yuan, L.; Craig, S.P.; McKerrow, J.H.; Wang, C.C.
J. Biol. Chem. 265, 13528-13532, 1990
A;Title: The hypoxanthine-guanine phosphoribosyltransferase of Schistosoma mansoni. Furt
A;Reference number: A37114; MUID:90337955; PMID:2199439
A;Accession: A37114
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-5 <YUA>
C;Keywords: glycosyltransferase; pentosyltransferase

```
  Query Match             28.1%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ISS 7
              :||
Db          1 MSS 3
```

RESULT 10
H44817
34.5K structural protein - Leuconostoc oenos phage P32 (fragment)
C;Species: Leuconostoc oenos phage P32
C;Date: 31-Mar-1993 #sequence_revision 22-May-1998 #text_change 22-May-1998
C;Accession: H44817
R;Arendt, E.K.; Lonvaud, A.; Hammes, W.P.
J. Gen. Microbiol. 137, 2135-2139, 1991
A;Title: Lysogeny in Leuconostoc oenos.
A;Reference number: A44817; MUID:92085033; PMID:1748868
A;Accession: H44817
A;Molecule type: protein
A;Residues: 1-5 <ARE>
A;Note: sequence extracted from NCBI backbone (NCBIP:70332)

```
  Query Match             28.1%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity   40.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 LTISS 7
              :  ||
Db          1 MATSS 5
```

RESULT 11
F44817
34.5K structural protein - Leuconostoc oenos phage P54 (fragment)
C;Species: Leuconostoc oenos phage P54
C;Date: 31-Mar-1993 #sequence_revision 22-May-1998 #text_change 22-May-1998
C;Accession: F44817
R;Arendt, E.K.; Lonvaud, A.; Hammes, W.P.
J. Gen. Microbiol. 137, 2135-2139, 1991
A;Title: Lysogeny in Leuconostoc oenos.
A;Reference number: A44817; MUID:92085033; PMID:1748868
A;Accession: F44817
A;Molecule type: protein
A;Residues: 1-5 <ARE>
A;Note: sequence extracted from NCBI backbone (NCBIP:70335)

```
  Query Match             28.1%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity   40.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 LTISS 7
```

```
              :  ||
Db          1 MATSS 5
```

RESULT 12
S69237
surface protein tetrabrachion heavy chain - Staphylothermus marinus (fragment)
C;Species: Staphylothermus marinus
C;Date: 04-Dec-1997 #sequence_revision 04-Dec-1997 #text_change 17-Mar-1999
C;Accession: S69237
R;Peters, J.; Nitsch, M.; Kuehlmorgen, B.; Golbik, R.; Lupas, A.; Kellermann, J.; E
J. Mol. Biol. 245, 385-401, 1995
A;Title: Tetrabrachion: a filamentous archaebacterial surface protein assembly of u
A;Reference number: S69237; MUID:95139068; PMID:7837271
A;Accession: S69237
A;Molecule type: protein
A;Residues: 1-5 <PET>
A;Experimental source: strain F1, DSM 3639
C;Keywords: cell wall; glycoprotein; heat-stable protein

```
  Query Match             28.1%;  Score 9;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TL 3
              ||
Db          2 TL 3
```

RESULT 13
A60986
N-formyl oligopeptide - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 31-Dec-1993
C;Accession: A60986
R;Broom, M.F.; Mellor, D.M.; Chadwick, V.S.
Experientia 45, 1097-1099, 1989
A;Title: Purification and amino acid sequencing of naturally occurring N-formyl-met
A;Reference number: A60986; MUID:90092408; PMID:2689204
A;Accession: A60986
A;Molecule type: protein
A;Residues: 1-6 <BRO>
C;Comment: This hexapeptide was the longest of several N-formyl oligopeptides report
F;1/Modified site: N-formylmethionine #status experimental

```
  Query Match             28.1%;  Score 9;  DB 2;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          3 FIL 5
```

RESULT 14
S14159
parasporal crystal protein CryIC - Bacillus thuringiensis (fragment)
N;Alternate names: delta-endotoxin
C;Species: Bacillus thuringiensis
C;Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 18-Jul-1997
C;Accession: S14159
R;Convents, D.; Cherlet, M.; van Damme, J.; Lasters, I.; Lauwereys, M.
Eur. J. Biochem. 195, 631-635, 1991
A;Title: Two structural domains as a general fold of the toxic fragment of the Bacil
A;Reference number: S14087; MUID:91153300; PMID:1847865
A;Accession: S14159
A;Molecule type: protein
A;Residues: 1-6 <CON>

```
  Query Match             28.1%;  Score 9;  DB 2;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

B44817
34.5K structural protein - Leuconostoc oenos phage PZt11-15 (fragment)
C;Species: Leuconostoc oenos phage PZt11-15
C;Date: 31-Mar-1993 #sequence_revision 22-May-1998 #text_change 22-May-1998
C;Accession: B44817
R;Arendt, E.K.; Lonvaud, A.; Hammes, W.P.
J. Gen. Microbiol. 137, 2135-2139, 1991
A;Title: Lysogeny in Leuconostoc oenos.
A;Reference number: A44817; MUID:92085033; PMID:1748868
A;Accession: B44817
A;Molecule type: protein
A;Residues: 1-5 <ARE>
A;Note: sequence extracted from NCBI backbone (NCBIP:70342)

Query Match 34.4%; Score 11; DB 2; Length 5;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LTISS 7
| ||
Db 1 LATSS 5

RESULT 4
D44817
35K structural protein - Leuconostoc oenos phage PAt5-12 (fragment)
C;Species: Leuconostoc oenos phage PAt5-12
C;Date: 31-Mar-1993 #sequence_revision 22-May-1998 #text_change 22-May-1998
C;Accession: D44817
R;Arendt, E.K.; Lonvaud, A.; Hammes, W.P.
J. Gen. Microbiol. 137, 2135-2139, 1991
A;Title: Lysogeny in Leuconostoc oenos.
A;Reference number: A44817; MUID:92085033; PMID:1748868
A;Accession: D44817
A;Molecule type: protein
A;Residues: 1-5 <ARE>
A;Note: sequence extracted from NCBI backbone (NCBIP:70340)

Query Match 34.4%; Score 11; DB 2; Length 5;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LTISS 7
| ||
Db 1 LATSS 5

RESULT 5
E30608
Ig kappa chain V-III region (Gag) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 16-Aug-1996
C;Accession: E30608
R;Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; Solc
J. Immunol. 142, 3158-3163, 1989
A;Title: Structural and idiotypic characterization of the L chains of human IgM autoanti
A;Reference number: A30601; MUID:89215279; PMID:2496160
A;Accession: E30608
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <GON>
C;Keywords: heterotetramer; immunoglobulin

Query Match 31.2%; Score 10; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LTIS 6
|| |
Db 4 LTQS 7

RESULT 6
T13892
cytochrome-c oxidase (EC 1.9.3.1) chain I [imported] - river lamprey mitochondrion
C;Species: mitochondrion Lampetra fluviatilis (river lamprey)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: T13892
R;Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A;Title: The main features of the craniate mitochondrial DNA between the ND1 and t
A;Reference number: Z17775; MUID:97398704; PMID:9254918
A;Accession: T13892
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3 <DEL>
A;Cross-references: EMBL:Y09528; NID:g2340016; PIDN:CAA70721.1; PID:g4379123
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

Query Match 28.1%; Score 9; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TL 3
||
Db 2 TL 3

RESULT 7
E42364
flagellar protein fliR - Salmonella typhimurium (fragment)
C;Species: Salmonella typhimurium
C;Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 30-Sep-1993
C;Accession: E42364
R;Vogler, A.P.; Homma, M.; Irikura, V.M.; Macnab, R.M.
J. Bacteriol. 173, 3564-3572, 1991
A;Title: Salmonella typhimurium mutants defective in flagellar filament regrowth a
A;Reference number: A42364; MUID:91258342; PMID:1646201
A;Accession: E42364
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-5 <VOG>
A;Cross-references: GB:M62408

Query Match 28.1%; Score 9; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TL 3
||
Db 3 TL 4

RESULT 8
E60274
major protein antigen MPT63 - Mycobacterium tuberculosis (fragment)
C;Species: Mycobacterium tuberculosis
C;Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 30-Sep-1993
C;Accession: E60274
R;Nagai, S.; Wiker, H.G.; Harboe, M.; Kinomoto, M.
Infect. Immun. 59, 372-382, 1991
A;Title: Isolation and partial characterization of major protein antigens in the c
A;Reference number: A60274; MUID:91099989; PMID:1898899
A;Accession: E60274
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-5 <NAG>

Query Match 28.1%; Score 9; DB 2; Length 5;
Best Local Similarity 25.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLT 4

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:56:46 ; Search time 11.6667 Seconds
(without alignments)
57.715 Million cell updates/sec

Title: US-09-712-819C-1
Perfect score: 32
Sequence: 1 FTLTISS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 457

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	11	34.4	4	2	A40135	branched-chain-ami
2	11	34.4	5	2	PT0644	T-cell receptor be
3	11	34.4	5	2	B44817	34.5K structural p
4	11	34.4	5	2	D44817	35K structural pro
5	10	31.2	7	2	E30608	Ig kappa chain V-I
6	9	28.1	3	3	T13892	cytochrome-c oxida
7	9	28.1	5	2	E42364	flagellar protein
8	9	28.1	5	2	E60274	major protein anti
9	9	28.1	5	2	A37114	hypoxanthine phosp
10	9	28.1	5	2	H44817	34.5K structural p
11	9	28.1	5	2	F44817	34.5K structural p
12	9	28.1	5	2	S69237	surface protein te
13	9	28.1	6	2	A60986	N-formyl oligopept
14	9	28.1	6	2	S14159	parasporal crystal
15	9	28.1	6	2	A43766	28K ubiquitin-immu
16	9	28.1	6	2	I37263	Y protein - human
17	9	28.1	6	2	I65546	MHC H2-L antigen -
18	9	28.1	7	2	S25266	pilE protein - Esc
19	9	28.1	7	2	A25269	sex pheromone cAM3
20	9	28.1	7	2	A30812	sex pheromone cCF1
21	9	28.1	7	2	PS0254	18K protein 5507 -
22	9	28.1	7	2	PN0649	pullulanase (EC 3.
23	9	28.1	7	2	S09066	globulin IV alpha
24	8	25.0	3	3	PT0622	T-cell receptor be
25	8	25.0	4	2	PT0696	T-cell receptor be
26	8	25.0	4	2	PT0645	T-cell receptor be
27	8	25.0	4	2	PT0712	T-cell receptor be
28	8	25.0	4	2	PT0698	T-cell receptor be
29	8	25.0	4	2	PT0551	T-cell receptor be
30	8	25.0	4	2	PT0697	T-cell receptor be
31	8	25.0	5	2	A60521	glycogen phosphory
32	8	25.0	5	2	I40698	biotin B - Citroba
33	8	25.0	5	2	I39964	ribosomal protein
34	8	25.0	5	2	I39966	ribosomal protein
35	8	25.0	5	2	I39965	ribosomal protein
36	8	25.0	5	2	T10954	hypothetical prote
37	8	25.0	5	2	I50385	myosin light chain
38	8	25.0	5	2	PT0308	Ig heavy chain CRD
39	8	25.0	5	2	PT0596	T-cell receptor be
40	8	25.0	5	2	PT0610	T-cell receptor be
41	8	25.0	5	2	PT0597	T-cell receptor be
42	8	25.0	5	2	PT0600	T-cell receptor be
43	8	25.0	5	2	PT0729	T-cell receptor be
44	8	25.0	5	2	PT0624	T-cell receptor be
45	8	25.0	5	2	PT0686	T-cell receptor be

ALIGNMENTS

RESULT 1
A40135
branched-chain-amino-acid transaminase (EC 2.6.1.42), mitochondrial - rat (fragment
N;Alternate names: branched-chain-amino-acid aminotransferase, mitochondrial
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-May-1992 #sequence_revision 13-May-1992 #text_change 30-Sep-1993
C;Accession: A40135
R;Hutson, S.M.; Wallin, R.; Hall, T.R.
submitted to the Protein Sequence Database, March 1992
A;Reference number: A40135
A;Accession: A40135
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-4 <HUT>
C;Keywords: aminotransferase; mitochondrion

Query Match 34.4%; Score 11; DB 2; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          5 ISS 7
              :||
Db          1 VSS 3
```

RESULT 2
PT0644
T-cell receptor beta chain V-D-J region (111-1G) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0644
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regio
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0644
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 34.4%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FT 2
              ||
Db          4 FT 5
```

RESULT 3

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2004, 11:18:42 ; Search time 34.6667 Seconds
(without alignments)
56.387 Million cell updates/sec

Title: US-09-712-819C-6
Perfect score: 33
Sequence: 1 FTLKISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 49082

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	60.6	7	9	US-09-791-378-198	Sequence 198, App
2	20	60.6	7	9	US-09-998-909-7	Sequence 7, Appli
3	20	60.6	7	9	US-09-826-290-205	Sequence 205, App
4	20	60.6	7	9	US-09-826-290-370	Sequence 370, App
5	20	60.6	7	10	US-09-791-393-145	Sequence 145, App
6	20	60.6	7	10	US-09-791-389-145	Sequence 145, App
7	20	60.6	7	12	US-10-601-100-105	Sequence 105, App
8	20	60.6	7	16	US-10-264-309-131	Sequence 131, App
9	20	60.6	7	16	US-10-264-309-460	Sequence 460, App
10	18	54.5	7	9	US-09-996-288-164	Sequence 164, App
11	18	54.5	7	10	US-09-996-265-164	Sequence 164, App
12	18	54.5	7	14	US-10-162-497-78	Sequence 78, Appl
13	18	54.5	7	15	US-10-461-863-164	Sequence 164, App
14	17	51.5	6	14	US-10-105-930-39	Sequence 39, Appl
15	17	51.5	6	14	US-10-043-539-30	Sequence 30, Appl
16	17	51.5	6	14	US-10-020-354-101	Sequence 101, App
17	17	51.5	7	9	US-09-996-288-135	Sequence 135, App
18	17	51.5	7	10	US-09-996-265-135	Sequence 135, App
19	17	51.5	7	13	US-10-044-034-13	Sequence 13, Appl
20	17	51.5	7	15	US-10-461-863-135	Sequence 135, App
21	16	48.5	6	9	US-09-727-963A-88	Sequence 88, Appl
22	16	48.5	6	14	US-10-211-994-23	Sequence 23, Appl
23	16	48.5	6	14	US-10-020-354-87	Sequence 87, Appl
24	16	48.5	7	9	US-09-989-789-396	Sequence 396, App
25	16	48.5	7	9	US-09-989-789-1971	Sequence 1971, Ap
26	16	48.5	7	9	US-09-989-789-3503	Sequence 3503, Ap
27	16	48.5	7	9	US-09-989-789-3542	Sequence 3542, Ap
28	16	48.5	7	9	US-09-989-789-3544	Sequence 3544, Ap
29	16	48.5	7	9	US-09-796-848A-20	Sequence 20, Appl
30	16	48.5	7	9	US-09-844-508-43	Sequence 43, Appl
31	16	48.5	7	9	US-09-808-387-44	Sequence 44, Appl
32	16	48.5	7	9	US-09-996-288-35	Sequence 35, Appl
33	16	48.5	7	9	US-09-996-288-96	Sequence 96, Appl
34	16	48.5	7	9	US-09-996-288-131	Sequence 131, App
35	16	48.5	7	10	US-09-990-186-396	Sequence 396, App
36	16	48.5	7	10	US-09-990-186-1971	Sequence 1971, Ap
37	16	48.5	7	10	US-09-990-186-3503	Sequence 3503, Ap
38	16	48.5	7	10	US-09-990-186-3542	Sequence 3542, Ap
39	16	48.5	7	10	US-09-990-186-3544	Sequence 3544, Ap
40	16	48.5	7	10	US-09-996-265-35	Sequence 35, Appl
41	16	48.5	7	10	US-09-996-265-96	Sequence 96, Appl
42	16	48.5	7	10	US-09-996-265-131	Sequence 131, App
43	16	48.5	7	10	US-09-989-994-396	Sequence 396, App
44	16	48.5	7	10	US-09-989-994-1971	Sequence 1971, Ap
45	16	48.5	7	10	US-09-989-994-3503	Sequence 3503, Ap

ALIGNMENTS

RESULT 1
US-09-791-378-198
; Sequence 198, Application US/09791378
; Patent No. US20020142303A1
; GENERAL INFORMATION:
; APPLICANT: Parekh, Rajesh
; TITLE OF INVENTION: PROTEINS, GENES AND THEIR USE FOR DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: SCHIZOPHRENIA
; FILE REFERENCE: 9195-061-999
; CURRENT APPLICATION NUMBER: US/09/791,378
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/750,395
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 677
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 198
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-378-198

Query Match 60.6%; Score 20; DB 9; Length 7;
Best Local Similarity 42.9%; Pred. No. 1e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTLKISR 7
              :| ::||
Db          1 YTFELSR 7
```

RESULT 2
US-09-998-909-7
; Sequence 7, Application US/09998909
; Patent No. US20020164664A1
; GENERAL INFORMATION:
; APPLICANT: Hlavaty, John
; APPLICANT: Briggman, Joseph
; TITLE OF INVENTION: Detection and Treatment of Prostate Cancer

```
;  FILE REFERENCE: MTP-027
;  CURRENT APPLICATION NUMBER: US/09/998,909
;  CURRENT FILING DATE:  2001-11-30
;  PRIOR APPLICATION NUMBER: US 60/250,284
;  PRIOR FILING DATE: 2000-11-30
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-998-909-7

  Query Match             60.6%;  Score 20;  DB 9;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1e+06;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              :| ::||
Db          1 YTFELSR 7


RESULT 3
US-09-826-290-205
; Sequence 205, Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
;  APPLICANT: Durham, L.Kathryn
;  APPLICANT:  Friedman, David L.
;  APPLICANT:  Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT:  Kimmel, Lida H.
;  APPLICANT:  Parekh, Rajesh Bhikhu
;  APPLICANT:  Potter, David M.
;  APPLICANT:  Rohlff, Christian
;  APPLICANT:  Silber, B. Michael
;  APPLICANT:  Stiger, Thomas R.
;  APPLICANT:  Sunderland, P. Trey
;  APPLICANT:  Townsend, Robert Reid
;  APPLICANT:  White, Frost
;  APPLICANT:  Williams, Stephen A.
;  TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
;  TITLE OF INVENTION:  Uses Therefor, Including Diagnosis and Treatment of
;  TITLE OF INVENTION:  Alzheimer's Disease
;  FILE REFERENCE: 2572-1-001 N2
;  CURRENT APPLICATION NUMBER: US/09/826,290
;  CURRENT FILING DATE:  2001-04-30
;  PRIOR APPLICATION NUMBER: US 60/194,504
;  PRIOR FILING DATE: 2000-04-03
;  PRIOR APPLICATION NUMBER: US 60/253,647
;  PRIOR FILING DATE: 2000-11-28
;  NUMBER OF SEQ ID NOS: 492
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 205
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-826-290-205

  Query Match             60.6%;  Score 20;  DB 9;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              || : |:|
Db          1 FTFEYSR 7


RESULT 4
US-09-826-290-370
; Sequence 370, Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
;  APPLICANT: Durham, L.Kathryn
;  APPLICANT:  Friedman, David L.
;  APPLICANT:  Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT:  Kimmel, Lida H.
;  APPLICANT:  Parekh, Rajesh Bhikhu
;  APPLICANT:  Potter, David M.
;  APPLICANT:  Rohlff, Christian
;  APPLICANT:  Silber, B. Michael
;  APPLICANT:  Stiger, Thomas R.
;  APPLICANT:  Sunderland, P. Trey
;  APPLICANT:  Townsend, Robert Reid
;  APPLICANT:  White, Frost
;  APPLICANT:  Williams, Stephen A.
;  TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
;  TITLE OF INVENTION:  Uses Therefor, Including Diagnosis and Treatment of
;  TITLE OF INVENTION:  Alzheimer's Disease
;  FILE REFERENCE: 2572-1-001 N2
;  CURRENT APPLICATION NUMBER: US/09/826,290
;  CURRENT FILING DATE:  2001-04-30
;  PRIOR APPLICATION NUMBER: US 60/194,504
;  PRIOR FILING DATE: 2000-04-03
;  PRIOR APPLICATION NUMBER: US 60/253,647
;  PRIOR FILING DATE: 2000-11-28
;  NUMBER OF SEQ ID NOS: 492
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 370
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-826-290-370

  Query Match             60.6%;  Score 20;  DB 9;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1e+06;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              :| ::||
Db          1 YTFELSR 7


RESULT 5
US-09-791-393-145
; Sequence 145, Application US/09791393
; Publication No. US20030032200A1
; GENERAL INFORMATION:
;  APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT: Parekh, Rajesh Bhikhu
;  APPLICANT: Rohlff, Christian
;  TITLE OF INVENTION: Proteins, Genes and Their Use for
;  TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
;  TITLE OF INVENTION: and Unipolar Depression
;  FILE REFERENCE: 2543-1-001 N1
;  CURRENT APPLICATION NUMBER: US/09/791,393
;  CURRENT FILING DATE: 2002-01-02
;  EARLIER APPLICATION NUMBER: GB 0004412.3
;  EARLIER FILING DATE: 2000-02-24
;  EARLIER APPLICATION NUMBER: GB 0030050.9
;  EARLIER FILING DATE: 2000-12-08
;  EARLIER APPLICATION NUMBER: US 60/254,830
;  EARLIER FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 145
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-791-393-145

  Query Match             60.6%;  Score 20;  DB 10;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1e+06;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              :| ::||
```

```
Db          1 YTFELSR 7


RESULT 6
US-09-791-389-145
; Sequence 145, Application US/09791389
; Publication No. US20030032773A1
; GENERAL INFORMATION:
;  APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT:  Parekh, Rajesh Bhikhu
;  APPLICANT:  Rohlff, Christian
;  APPLICANT:  Terrett, Jonathan Alexander
;  APPLICANT:  Tyson, Kerry Louise
;  TITLE OF INVENTION: Proteins, Genes and Their Use for
;  TITLE OF INVENTION:  Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
;  TITLE OF INVENTION:  and Unipolar Depression
;  FILE REFERENCE: 2543-1-001 N2
;  CURRENT APPLICATION NUMBER: US/09/791,389
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: GB 0004412.3
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: GB 0030050.9
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: US 60/254,830
;  PRIOR FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 145
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-791-389-145

  Query Match             60.6%;  Score 20;  DB 10;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1e+06;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              :| ::||
Db          1 YTFELSR 7


RESULT 7
US-10-601-100-105
; Sequence 105, Application US/10601100
; Publication No. US20040072261A1
; GENERAL INFORMATION:
;  APPLICANT: INNOGENETICS N.V.
;  TITLE OF INVENTION: Method for the Diagnosis and Differential Diagnosis of
;  TITLE OF INVENTION:  Neurological Diseases
;  FILE REFERENCE: 11362.0038.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/601,100
;  CURRENT FILING DATE:  2003-06-20
;  PRIOR APPLICATION NUMBER: EP 02447121.1
;  PRIOR FILING DATE: 2002-06-21
;  PRIOR APPLICATION NUMBER: US 60/396,437
;  PRIOR FILING DATE: 2002-07-17
;  NUMBER OF SEQ ID NOS: 113
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 105
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-10-601-100-105

  Query Match             60.6%;  Score 20;  DB 12;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1e+06;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              :| ::||
Db          1 YTFELSR 7


RESULT 8
US-10-264-309-131
; Sequence 131, Application US/10264309
; Publication No. US20040022794A1
; GENERAL INFORMATION:
;  APPLICANT: DURHAM, L. KATHRYN
;  APPLICANT:  FRIEDMAN, DAVID L.
;  APPLICANT:  HERATH, HERATH
;  APPLICANT:  KIMMEL, LIDA H.
;  APPLICANT:  PAREKH, RAJESH B.
;  APPLICANT:  POTTER, DAVID M.
;  APPLICANT:  ROHLFF, CHRISTIAN
;  APPLICANT:  SILBER, B. MICHAEL
;  APPLICANT:  SNYDER, PETER J.
;  APPLICANT:  SOARES, HOLLY D.
;  APPLICANT:  STIGER, THOMAS R.
;  APPLICANT:  SUNDERLAND, P. TREY
;  APPLICANT:  TOWNSEND, ROBERT R.
;  APPLICANT:  WHITE, W. FROST
;  APPLICANT:  WILLIAMS, STEPHEN A.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES, POLYPEPTIDES AND USES THEREFOR,
;  TITLE OF INVENTION:  INCLUDING DIAGNOSIS AND TREATMENT OF ALZHEIMER'S DISEASE
;  FILE REFERENCE: POA-002.01
;  CURRENT APPLICATION NUMBER: US/10/264,309
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: 60/326,708
;  PRIOR FILING DATE: 2001-10-03
;  NUMBER OF SEQ ID NOS: 491
;  SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 131
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-264-309-131

  Query Match             60.6%;  Score 20;  DB 16;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              || : ||
Db          1 FTFEYSR 7


RESULT 9
US-10-264-309-460
; Sequence 460, Application US/10264309
; Publication No. US20040022794A1
; GENERAL INFORMATION:
;  APPLICANT: DURHAM, L. KATHRYN
;  APPLICANT:  FRIEDMAN, DAVID L.
;  APPLICANT:  HERATH, HERATH
;  APPLICANT:  KIMMEL, LIDA H.
;  APPLICANT:  PAREKH, RAJESH B.
;  APPLICANT:  POTTER, DAVID M.
;  APPLICANT:  ROHLFF, CHRISTIAN
;  APPLICANT:  SILBER, B. MICHAEL
;  APPLICANT:  SNYDER, PETER J.
;  APPLICANT:  SOARES, HOLLY D.
;  APPLICANT:  STIGER, THOMAS R.
;  APPLICANT:  SUNDERLAND, P. TREY
;  APPLICANT:  TOWNSEND, ROBERT R.
;  APPLICANT:  WHITE, W. FROST
;  APPLICANT:  WILLIAMS, STEPHEN A.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES, POLYPEPTIDES AND USES THEREFOR,
;  TITLE OF INVENTION:  INCLUDING DIAGNOSIS AND TREATMENT OF ALZHEIMER'S DISEASE
;  FILE REFERENCE: POA-002.01
;  CURRENT APPLICATION NUMBER: US/10/264,309
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: 60/326,708
;  PRIOR FILING DATE: 2001-10-03
;  NUMBER OF SEQ ID NOS: 491
```

```
;  SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 460
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-264-309-460

  Query Match             60.6%;  Score 20;  DB 16;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1e+06;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              :| ::||
Db          1 YTFELSR 7


RESULT 10
US-09-996-288-164
; Sequence 164, Application US/09996288
; Patent No. US20020177126A1
; GENERAL INFORMATION:
;  APPLICANT: Young, James
;  APPLICANT:  Scott, Koenig
;  APPLICANT:  Leslie, Johnson
;  TITLE OF INVENTION: Methods of Adminstering/Dosing Anti-RSV Antibodies for Prophylaxi
;  TITLE OF INVENTION:  and Treatment
;  FILE REFERENCE: 10271-047-999
;  CURRENT APPLICATION NUMBER: US/09/996,288
;  CURRENT FILING DATE:  2001-11-28
;  NUMBER OF SEQ ID NOS: 259
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-996-288-164

  Query Match             54.5%;  Score 18;  DB 9;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLKIS 6
              |:|:|
Db          2 TMKLS 6


RESULT 11
US-09-996-265-164
; Sequence 164, Application US/09996265
; Publication No. US20030091584A1
; GENERAL INFORMATION:
;  APPLICANT: Young, James
;  APPLICANT:  Scott, Koenig
;  APPLICANT:  Leslie, Johnson
;  TITLE OF INVENTION: Methods of Adminstering/Dosing Anti-RSV Antibodies for Prophylaxi
;  TITLE OF INVENTION:  and Treatment
;  FILE REFERENCE: 10271-048-999
;  CURRENT APPLICATION NUMBER: US/09/996,265
;  CURRENT FILING DATE:  2001-11-28
;  NUMBER OF SEQ ID NOS: 259
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-996-265-164

  Query Match             54.5%;  Score 18;  DB 10;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLKIS 6
              |:|:|
Db          2 TMKLS 6


RESULT 12
US-10-162-497-78
; Sequence 78, Application US/10162497
; Publication No. US20030158398A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, H.
;  APPLICANT:  Freimer, N.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING CHROMOSOME-18p RELATED DISORDERS
;  FILE REFERENCE: 7853-138
;  CURRENT APPLICATION NUMBER: US/10/162,497
;  CURRENT FILING DATE:  2002-06-04
;  PRIOR APPLICATION NUMBER: US/09/657,474
;  PRIOR FILING DATE: 2000-09-07
;  PRIOR APPLICATION NUMBER: 09/268,992
;  PRIOR FILING DATE: 1999-03-16
;  PRIOR APPLICATION NUMBER: 09/236,134
;  PRIOR FILING DATE: 1999-01-22
;  PRIOR APPLICATION NUMBER: 60/106,056
;  PRIOR FILING DATE: 1998-10-28
;  PRIOR APPLICATION NUMBER: 60/088,312
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/078,044
;  PRIOR FILING DATE: 1998-03-16
;  NUMBER OF SEQ ID NOS: 84
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 78
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-162-497-78

  Query Match             54.5%;  Score 18;  DB 14;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKIS 6
              :|:| |
Db          2 YTMKYS 7


RESULT 13
US-10-461-863-164
; Sequence 164, Application US/10461863
; Publication No. US20040018200A1
; GENERAL INFORMATION:
;  APPLICANT: Oliver, Cynthia
;  APPLICANT:  Allan, Christian
;  APPLICANT:  Chang, Stephen
;  TITLE OF INVENTION: STABILIZED ANTI-RESPIRATORY SYNCYTIAL VIRUS (RSV) ANTIBODY FORM
;  FILE REFERENCE: 10271-071-999
;  CURRENT APPLICATION NUMBER: US/10/461,863
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: 60/388,920
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 209
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-461-863-164

  Query Match             54.5%;  Score 18;  DB 15;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLKIS 6
              |:|:|
Db          2 TMKLS 6
```

```
RESULT 14
US-10-105-930-39
; Sequence 39, Application US/10105930
; Publication No. US20030009018A1
; GENERAL INFORMATION:
;  APPLICANT: Maeda, Masatsugu
;  APPLICANT:  Yaguchi, No. US20030009018A1iko
;  TITLE OF INVENTION: NOVEL HEMOPOIETIN RECEPTOR PROTEIN, NR12
;  FILE REFERENCE: 06501-105US1
;  CURRENT APPLICATION NUMBER: US/10/105,930
;  CURRENT FILING DATE:  2002-03-25
;  PRIOR APPLICATION NUMBER: PCT/JP00/06654
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: JP 2000-240397
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: JP 11-273358
;  PRIOR FILING DATE: 1999-09-27
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-105-930-39

  Query Match           51.5%;  Score 17;  DB 14;  Length 6;
  Best Local Similarity  60.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLKI 5
              :||:|
Db          1 YTLQI 5


RESULT 15
US-10-043-539-30
; Sequence 30, Application US/10043539
; Publication No. US20030114650A1
; GENERAL INFORMATION:
;  APPLICANT: Cheung, Ambrose
;  APPLICANT:  Manna, Adar
;  APPLICANT:  Zhang, Gongyi
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AFFECTING VIRULENCE DETERMINANTS IN
;  TITLE OF INVENTION:  BACTERIA
;  FILE REFERENCE: DC-0199
;  CURRENT APPLICATION NUMBER: US/10/043,539
;  CURRENT FILING DATE:  2002-01-11
;  PRIOR APPLICATION NUMBER: US 60/261,233
;  PRIOR FILING DATE: 2001-01-12
;  PRIOR APPLICATION NUMBER: US 60/261,607
;  PRIOR FILING DATE: 2001-01-12
;  PRIOR APPLICATION NUMBER: US 60/289,601
;  PRIOR FILING DATE: 2001-05-08
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-10-043-539-30

  Query Match           51.5%;  Score 17;  DB 14;  Length 6;
  Best Local Similarity  60.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LKISR 7
              :||:|
Db          2 IKITR 6


Search completed: June  1, 2004, 11:31:24
```

```
Job time : 35.6667 secs
```

GenCore version 5.1.6

CM protein - protein search, using sw model

Run on: June 1, 2004, 11:13:51 ; Search time 13.6667 Seconds
(without alignments)
26.443 Million cell updates/sec

Title: US-09-712-819C-6
Perfect score: 33
Sequence: 1 FTLKISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 57228

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	60.6	7	1	US-08-222-619-8	Sequence 8, Appli
2	20	60.6	7	5	PCT-US95-04075-8	Sequence 8, Appli
3	19	57.6	6	6	5252328-8	Patent No. 5252328
4	18	54.5	7	1	US-08-346-333-24	Sequence 24, Appl
5	18	54.5	7	4	US-09-268-992-78	Sequence 78, Appl
6	18	54.5	7	4	US-09-657-474-78	Sequence 78, Appl
7	18	54.5	7	4	US-08-753-750B-20	Sequence 20, Appl
8	18	54.5	7	5	PCT-US91-07506-24	Sequence 24, Appl
9	17	51.5	7	2	US-09-174-060-17	Sequence 17, Appl
10	17	51.5	7	3	US-08-654-623-69	Sequence 69, Appl
11	17	51.5	7	3	US-08-338-382-17	Sequence 17, Appl
12	16	48.5	5	3	US-08-591-632-23	Sequence 23, Appl
13	16	48.5	5	4	US-09-611-451-23	Sequence 23, Appl
14	16	48.5	6	1	US-07-717-331F-5	Sequence 5, Appli
15	16	48.5	6	4	US-08-757-425B-44	Sequence 44, Appl
16	16	48.5	7	1	US-07-634-641-12	Sequence 12, Appl
17	16	48.5	7	1	US-08-136-743B-55	Sequence 55, Appl
18	16	48.5	7	3	US-09-040-216-28	Sequence 28, Appl
19	16	48.5	7	3	US-09-173-941-72	Sequence 72, Appl
20	16	48.5	7	4	US-09-423-468A-6	Sequence 6, Appli
21	16	48.5	7	4	US-09-756-223A-26	Sequence 26, Appl
22	16	48.5	7	4	US-09-989-789-396	Sequence 396, App
23	16	48.5	7	4	US-09-989-789-1971	Sequence 1971, Ap
24	16	48.5	7	4	US-09-989-789-3503	Sequence 3503, Ap
25	16	48.5	7	4	US-09-989-789-3542	Sequence 3542, Ap
26	16	48.5	7	4	US-09-989-789-3544	Sequence 3544, Ap
27	16	48.5	7	4	US-09-494-190-72	Sequence 72, Appl
28	15	45.5	4	1	US-08-305-871A-17	Sequence 17, Appl
29	15	45.5	4	4	US-08-788-822A-1	Sequence 1, Appli
30	15	45.5	4	4	US-08-134-231C-36	Sequence 36, Appl
31	15	45.5	4	4	US-08-728-160-36	Sequence 36, Appl
32	15	45.5	5	1	US-08-136-743B-63	Sequence 63, Appl
33	15	45.5	5	3	US-09-040-216-55	Sequence 55, Appl
34	15	45.5	5	3	US-08-591-632-17	Sequence 17, Appl
35	15	45.5	5	3	US-08-591-632-26	Sequence 26, Appl
36	15	45.5	5	4	US-09-611-451-17	Sequence 17, Appl
37	15	45.5	5	4	US-09-611-451-26	Sequence 26, Appl
38	15	45.5	6	1	US-08-136-743B-62	Sequence 62, Appl
39	15	45.5	6	1	US-08-357-264-6	Sequence 6, Appli
40	15	45.5	6	1	US-08-297-731-4	Sequence 4, Appli
41	15	45.5	6	1	US-08-222-619-21	Sequence 21, Appl
42	15	45.5	6	1	US-08-290-448A-41	Sequence 41, Appl
43	15	45.5	6	1	US-08-672-514-6	Sequence 6, Appli
44	15	45.5	6	1	US-08-290-448A-41	Sequence 41, Appl
45	15	45.5	6	1	US-08-175-069A-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-08-222-619-8
; Sequence 8, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-222-619-8

Query Match 60.6%; Score 20; DB 1; Length 7;
Best Local Similarity 57.1%; Pred. No. 3e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FTLKISR 7
|| : ||
Db 1 FTFEYSR 7

RESULT 2
PCT-US95-04075-8
; Sequence 8, Application PC/TUS9504075
; GENERAL INFORMATION:

```
;    APPLICANT:  AMGEN INC.
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04075
;      FILING DATE:
;      CLASSIFICATION:
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US95-04075-8

  Query Match             60.6%;  Score 20;  DB 5;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 3e+05;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLKISR 7
              || : |
Db          1 FTFEYSR 7


RESULT 3
5252328-8
;Patent No. 5252328
;    APPLICANT: FAULDS,DARYL;VISHOOT,MIMI;BROOKS,EMILY
;    TITLE OF INVENTION: MYCOPLASMA HYOPNEUMONIAE ANTIGEN AND USES
;THEREFORE
;    NUMBER OF SEQUENCES: 15
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/335,726
;      FILING DATE: 07-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 30,130
;      FILING DATE: 26-MAR-1987
;SEQ ID NO:8:
;      LENGTH: 6
5252328-8

  Query Match             57.6%;  Score 19;  DB 6;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 3e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKI 5
              | |||
Db          2 FVLKI 6


RESULT 4
US-08-346-333-24
; Sequence 24, Application US/08346333
; Patent No. 5677153
;  GENERAL INFORMATION:
;    APPLICANT:  Botstein, David
;    APPLICANT:  Palzkill, Timothy
;    TITLE OF INVENTION:  Methods for modifying DNA and for
;    TITLE OF INVENTION:  detecting effects of such modification on interaction of
;    TITLE OF INVENTION:  encoded modified polypeptides with target substrates.
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard F. Trecartin
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/346,333
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/039,501
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/602,158
;      FILING DATE:  22-OCT-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-53469/RFT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-346-333-24

  Query Match             54.5%;  Score 18;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLKI 5
              ||||
Db          2 TLKI 5


RESULT 5
US-09-268-992-78
; Sequence 78, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
;  APPLICANT: Chen, H.
;  APPLICANT: Freimer, N.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
;  FILE REFERENCE: 7853-138
;  CURRENT APPLICATION NUMBER: US/09/268,992
;  CURRENT FILING DATE: 1999-03-16
;  EARLIER APPLICATION NUMBER: 09/236,134
;  EARLIER FILING DATE: 1999-01-22
;  EARLIER APPLICATION NUMBER: 60/106,056
;  EARLIER FILING DATE: 1998-10-28
;  EARLIER APPLICATION NUMBER: 60/088,312
;  EARLIER FILING DATE: 1998-06-05
;  EARLIER APPLICATION NUMBER: 60/078,044
;  EARLIER FILING DATE: 1998-03-16
;  NUMBER OF SEQ ID NOS: 84
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 78
;   LENGTH: 7
;   TYPE: PRT
```

```
;   ORGANISM: Homo sapiens
US-09-268-992-78

  Query Match             54.5%;  Score 18;  DB 4;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKIS 6
              :|:| |
Db          2 YTMKYS 7


RESULT 6
US-09-657-474-78
; Sequence 78, Application US/09657474
; Patent No. 6399762
; GENERAL INFORMATION:
;  APPLICANT: Chen, H.
;  APPLICANT:  Freimer, N.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING CHROMOSOME-18p RELATED DISORDERS
;  FILE REFERENCE: 7853-138
;  CURRENT APPLICATION NUMBER: US/09/657,474
;  CURRENT FILING DATE:  2000-09-07
;  PRIOR APPLICATION NUMBER: 09/268,992
;  PRIOR FILING DATE: 1999-03-16
;  PRIOR APPLICATION NUMBER: 09/236,134
;  PRIOR FILING DATE: 1999-01-22
;  PRIOR APPLICATION NUMBER: 60/106,056
;  PRIOR FILING DATE: 1998-10-28
;  PRIOR APPLICATION NUMBER: 60/088,312
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/078,044
;  PRIOR FILING DATE: 1998-03-16
;  NUMBER OF SEQ ID NOS: 84
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 78
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-657-474-78

  Query Match             54.5%;  Score 18;  DB 4;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKIS 6
              :|:| |
Db          2 YTMKYS 7


RESULT 7
US-08-753-750B-20
; Sequence 20, Application US/08753750B
; Patent No. 6610506
; GENERAL INFORMATION:
;  APPLICANT: Lo, Reggie Y.C.
;  APPLICANT:  Schryvers, Anthony B.
;  APPLICANT:  Potter, Andrew A.
;  TITLE OF INVENTION: TRANSFERRIN BINDING PROTEINS OF
;  TITLE OF INVENTION:  PASTEURELLA HAEMOLYTICA AND VACCINES CONTAINING THE SAME
;  FILE REFERENCE: A34762 021645.0105
;  CURRENT APPLICATION NUMBER: US/08/753,750B
;  CURRENT FILING DATE:  1996-11-29
;  PRIOR APPLICATION NUMBER: CA 2,164,274
;  PRIOR FILING DATE: 1995-12-01
;  PRIOR APPLICATION NUMBER: 60/008,569
;  PRIOR FILING DATE: 1995-12-01
;  NUMBER OF SEQ ID NOS: 68
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-08-753-750B-20

  Query Match             54.5%;  Score 18;  DB 4;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLKI 5
              ||| :
Db          1 FTLSV 5


RESULT 8
PCT-US91-07506-24
; Sequence 24, Application PC/TUS9107506
;  GENERAL INFORMATION:
;    APPLICANT:  Botstein, David
;    APPLICANT:  Palzkill, Timothy
;    TITLE OF INVENTION:  Methods for modifying DNA and for
;    TITLE OF INVENTION:  detecting effects of such modification on interaction of
;    TITLE OF INVENTION:  encoded modified polypeptides with target substrates.
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard F. Trecartin
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/07506
;      FILING DATE:  19911021
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  FP-53469-PC/RFT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US91-07506-24

  Query Match             54.5%;  Score 18;  DB 5;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLKI 5
              ||||
Db          2 TLKI 5


RESULT 9
US-09-174-060-17
; Sequence 17, Application US/09174060
; Patent No. 5989554
;  GENERAL INFORMATION:
;    APPLICANT:  Knuth, Mark W
;    APPLICANT:  Haak-Frendscho, Mary
;    APPLICANT:  Shultz, John W
;    APPLICANT:  Lesley, Scott A
;    APPLICANT:  Villars, Catherine E
```

```
;    TITLE OF INVENTION:  HIGH LEVEL EXPRESSION AND FACILE
;    TITLE OF INVENTION:  PURIFICATION OF PROTEINS, PEPTIDES AND CONJUGATES FOR
;    TITLE OF INVENTION:  IMMUNIZATION, PURIFICATION AND DETECTION APPLICATIONS
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ross & Stevens, S.C.
;      STREET:  1 South Pinckney St.
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  USA
;      ZIP:  53701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/174,060
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/338,382
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sara, Charles S
;      REGISTRATION NUMBER:  30,492
;      REFERENCE/DOCKET NUMBER:  34506.024
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-257-5353
;      TELEFAX:  608-257-9175
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-174-060-17

  Query Match           51.5%;  Score 17;  DB 2;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLK 4
              :|||
Db          1 YTLK 4


RESULT 10
US-08-654-623-69
; Sequence 69, Application US/08654623
; Patent No. 6010884
;  GENERAL INFORMATION:
;    APPLICANT:  Griffiths, Andrew D
;    APPLICANT:  Holliger, Kaspar-Philipp
;    APPLICANT:  Nissim, Ahuva
;    APPLICANT:  Fisch, Igor
;    APPLICANT:  Winter, Gregory P
;    TITLE OF INVENTION:  Recombinant Binding Proteins and Peptides
;    NUMBER OF SEQUENCES:  71
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/654,623
;      FILING DATE:  29-MAY-1996
;      CLASSIFICATION:  435
;      CLASSIFICATION:  (C12N 1/21, C12R 1:19)
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9225453.1
;      FILING DATE:  04-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9300816.7
;      FILING DATE:  16-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 93303614.7
;      FILING DATE:  10-MAY-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9319969.3
;      FILING DATE:  22-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/02492
;      FILING DATE:  03-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9412147.2
;      FILING DATE:  17-JUN-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB94/02662
;      FILING DATE:  05-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/448,418
;      FILING DATE:  02-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  David W. Clough
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  28111/33259
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-654-623-69

  Query Match           51.5%;  Score 17;  DB 3;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLKIS 6
              :||:|
Db          1 SLKVS 5


RESULT 11
US-08-338-382-17
; Sequence 17, Application US/08338382
; Patent No. 6069230
;  GENERAL INFORMATION:
;    APPLICANT:  Knuth, Mark W
;    APPLICANT:  Haak-Frendscho, Mary
;    APPLICANT:  Shultz, John W
;    APPLICANT:  Lesley, Scott A
;    APPLICANT:  Villars, Catherine E
;    TITLE OF INVENTION:  HIGH LEVEL EXPRESSION AND FACILE
;    TITLE OF INVENTION:  PURIFICATION OF PROTEINS, PEPTIDES AND CONJUGATES FOR
;    TITLE OF INVENTION:  IMMUNIZATION, PURIFICATION AND DETECTION APPLICATIONS
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ross & Stevens, S.C.
;      STREET:  1 South Pinckney St.
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  USA
;      ZIP:  53701
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/338,382
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sara, Charles S
;       REGISTRATION NUMBER:  30,492
;       REFERENCE/DOCKET NUMBER:  34506.024
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  608-257-5353
;       TELEFAX:  608-257-9175
;   INFORMATION FOR SEQ ID NO:  17:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-338-382-17

  Query Match             51.5%;  Score 17;  DB 3;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLK 4
              :|||
Db          1 YTLK 4


RESULT 12
US-08-591-632-23
; Sequence 23, Application US/08591632
; Patent No. 6261558
;  GENERAL INFORMATION:
;    APPLICANT:  Barbas, Carlos F.
;    APPLICANT:  Burton, Dennis R.
;    APPLICANT:  Lerner, Righard A.
;    TITLE OF INVENTION:  SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
;    TITLE OF INVENTION:  ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Scripps Research Institute, Office of
;      ADDRESSEE:  Patent Counsel
;      STREET:  10550 No. 6261558th Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/591,632
;      FILING DATE:  19-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/11907
;      FILING DATE:  19-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/308,841
;      FILING DATE:  19-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/233,619
;      FILING DATE:  26-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/139,409
;      FILING DATE:  19-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  TSRI 332.3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 784-2937
;      TELEFAX:  (619) 784-9399
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-591-632-23

  Query Match             48.5%;  Score 16;  DB 3;  Length 5;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLK 4
              |||:
Db          2 FTLQ 5


RESULT 13
US-09-611-451-23
; Sequence 23, Application US/09611451
; Patent No. 6395275
;  GENERAL INFORMATION:
;       APPLICANT: Barbas, Carlos F.
;                  Burton, Dennis R.
;                  Lerner, Righard A.
;       TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
;                           ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
;       NUMBER OF SEQUENCES: 92
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: The Scripps Research Institute, Office of
;                       Patent Counsel
;            STREET: 10550 No. 6395275th Torrey Pines Road, TPC 8
;            CITY: La Jolla
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 92037
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/611,451
;            FILING DATE: 06-Jul-2000
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/591,632
;            FILING DATE: 2001-10-29
;            APPLICATION NUMBER: US 08/308,841
;            FILING DATE: 19-SEP-1994
;            APPLICATION NUMBER: US 08/233,619
;            FILING DATE: 26-APR-1994
;            APPLICATION NUMBER: US 08/139,409
;            FILING DATE: 19-OCT-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Fitting, Thomas
;            REGISTRATION NUMBER: 34,163
;            REFERENCE/DOCKET NUMBER: TSRI 332.3
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (619) 784-2937
;            TELEFAX: (619) 784-9399
;  INFORMATION FOR SEQ ID NO: 23:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 5 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 23:
```

US-09-611-451-23

Query Match 48.5%; Score 16; DB 4; Length 5;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FTLK 4
              |||:
Db          2 FTLQ 5
```

RESULT 14
US-07-717-331F-5
; Sequence 5, Application US/07717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatability Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/717,331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-717-331F-5

Query Match 48.5%; Score 16; DB 1; Length 6;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 LKIS 6
              ||:|
Db          2 LKVS 5
```

RESULT 15
US-08-757-425B-44
; Sequence 44, Application US/08757425B
; Patent No. 6500660
; GENERAL INFORMATION:
; APPLICANT: Fastrez, Jacques
; TITLE OF INVENTION: Chimeric Target Molecules Having A Regulatable Activity
; FILE REFERENCE: 100390-09640
; CURRENT APPLICATION NUMBER: US/08/757,425B
; CURRENT FILING DATE: 1996-11-27
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mimetope
US-08-757-425B-44

Query Match 48.5%; Score 16; DB 4; Length 6;
Best Local Similarity 60.0%; Pred. No. 3e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 TLKIS 6
              | |:|
Db          2 TAKVS 6
```

Search completed: June 1, 2004, 11:20:18
Job time : 13.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:54:42 ; Search time 13 Seconds
(without alignments)
51.795 Million cell updates/sec

Title: US-09-712-819D-12
Perfect score: 33
Sequence: 1 FTLEISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 457

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	13	39.4	6	2	B26206	alpha-1,4-glucan-p
2	13	39.4	7	2	E48394	glycoprotein compo
3	13	39.4	7	2	I48086	DNA topoisomerase
4	13	39.4	7	2	B48394	major fat-globule
5	12	36.4	5	2	T14910	hypothetical prote
6	11	33.3	4	2	A61300	22K superhelical D
7	11	33.3	5	2	PT0644	T-cell receptor be
8	11	33.3	7	2	S19630	ribosomal protein
9	10	30.3	5	2	A60521	glycogen phosphory
10	10	30.3	7	2	B39127	phosphotransferase
11	10	30.3	7	2	PS0254	18K protein 5507 -
12	10	30.3	7	2	S33244	neuromodulatory pe
13	10	30.3	7	2	S33245	neuromodulatory pe
14	9	27.3	3	3	T13892	cytochrome-c oxida
15	9	27.3	5	2	E42364	flagellar protein
16	9	27.3	5	2	A44955	alkanal monooxygen
17	9	27.3	5	2	S11127	phosphoprotein, bo
18	9	27.3	5	2	PT0525	T-cell receptor be
19	9	27.3	5	2	PT0577	T-cell receptor be
20	9	27.3	5	2	PT0565	T-cell receptor be
21	9	27.3	5	2	PT0700	T-cell receptor be
22	9	27.3	5	2	S69237	surface protein te
23	9	27.3	6	2	B34835	dnaA protein - Pse
24	9	27.3	6	2	A60986	N-formyl oligopept
25	9	27.3	6	2	T11779	phosphoglycerate t
26	9	27.3	6	2	A43766	28K ubiquitin-immu
27	9	27.3	6	2	JH0784	neuropeptide TE-6
28	9	27.3	6	2	I37263	Y protein - human
29	9	27.3	6	2	S78764	ribosomal protein
30	9	27.3	6	2	A20186	fatty-acid synthas
31	9	27.3	6	2	I65546	MHC H2-L antigen -
32	9	27.3	6	2	PT0518	T-cell receptor be
33	9	27.3	6	2	PT0662	T-cell receptor be
34	9	27.3	7	2	JN0859	peptidyl-dipeptida
35	9	27.3	7	2	A15398	choline oxidase (E
36	9	27.3	7	2	S25266	pilE protein - Esc
37	9	27.3	7	2	A25269	sex pheromone cAM3
38	9	27.3	7	2	A30812	sex pheromone cCF1
39	9	27.3	7	2	A28709	phosphonoacetaldeh
40	9	27.3	7	2	PN0150	omega-gliadine 1'
41	9	27.3	7	2	PT0269	Ig heavy chain CRD
42	9	27.3	7	2	B30608	Ig kappa chain V-I
43	9	27.3	7	2	PT0671	T-cell receptor be
44	9	27.3	7	2	A38671	peptidylglycine mo
45	9	27.3	7	2	A58718	carnocin UI49 - Ca

ALIGNMENTS

RESULT 1
B26206
alpha-1,4-glucan-protein synthase (UDP-forming) (EC 2.4.1.112) - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Oct-1992 #sequence_revision 30-Oct-1992 #text_change 13-Sep-1996
C;Accession: B26206; A26206
R;Larner, J.; Sanger, F.
J. Mol. Biol. 11, 491-500, 1965
A;Title: The amino acid sequence of the phosphorylation site of muscle uridine diphospho
A;Reference number: A26206
A;Accession: B26206
A;Molecule type: protein
A;Residues: 1-6 <LAR>
A;Experimental source: muscle
A;Note: Lys-1 was also found
C;Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein
F;4/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 39.4%; Score 13; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          4 EIS 6
              |||
Db          2 EIS 4
```

RESULT 2
E48394
glycoprotein component 16/major fat-globule membrane protein/MFG-E8 homolog - bovine (fra
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 07-Feb-1997
C;Accession: E48394
R;Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A;Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.
A;Reference number: A48394; MUID:93250576; PMID:8485470
A;Accession: E48394
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MAT>
A;Experimental source: milk
A;Note: sequence extracted from NCBI backbone (NCBIP:131450)
C;Keywords: glycoprotein

Query Match 39.4%; Score 13; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          4 EISR 7
              |::|
```

```
Db          4 ELAR 7


RESULT 3
I48086
DNA topoisomerase II alpha - Chinese hamster (fragment)
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I48086
R;Ng, S.W.; Eder, J.P.; Schnipper, L.E.; Chan, V.T.W.
J. Biol. Chem. 270, 25850-25858, 1995
A;Title: Molecular cloning and characterization of the promoter for the Chinese hamster
A;Reference number: I48086; MUID:96029684; PMID:7592770
A;Accession: I48086
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-7 <RES>
A;Cross-references: EMBL:U34196; NID:g1041231; PIDN:AAC52315.1; PID:g1041232

  Query Match             39.4%;  Score 13;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LEIS 6
              :|:|
Db          1 MELS 4


RESULT 4
B48394
major fat-globule membrane protein GP 55 - guinea pig (fragment)
C;Species: Cavia porcellus (guinea pig)
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995
C;Accession: B48394
R;Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A;Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.
A;Reference number: A48394; MUID:93250576; PMID:8485470
A;Accession: B48394
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MAT>
A;Experimental source: milk
A;Note: sequence extracted from NCBI backbone (NCBIP:131444)

  Query Match             39.4%;  Score 13;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EISR 7
              |::|
Db          4 ELAR 7


RESULT 5
T14910
hypothetical protein - parsley
C;Species: Petroselinum crispum (parsley)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14910
R;Kircher, S.; Ledger, S.; Hayashi, H.; Weisshaar, B.; Schafer, E.; Frohnmeyer, H.
Mol. Gen. Genet. 257, 595-605, 1998
A;Title: CPRF4a, a novel plant bZIP protein of the CPRF family: comparative analysis of
A;Reference number: Z18261; MUID:98265918; PMID:9604882
A;Accession: T14910
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-5 <KIR>
A;Cross-references: EMBL:Y10810; NID:g3336904; PIDN:CAA71769.1; PID:g3336905
A;Experimental source: ssp. Hamburger Schnitt

  Query Match             36.4%;  Score 12;  DB 2;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ISR 7
              :||
Db          2 VSR 4


RESULT 6
A61300
22K superhelical DNA-binding protein - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
C;Accession: A61300
R;Kishi, F.; Ebina, Y.; Miki, T.; Nakazawa, T.; Nakazawa, A.
J. Biochem. 92, 1059-1068, 1982
A;Title: Purification and characterization of a protein from Escherichia coli which forr
A;Reference number: A61300; MUID:83082696; PMID:6294066
A;Accession: A61300
A;Molecule type: protein
A;Residues: 1-4 <KIS>
C;Comment: This protein resembles some of the histone-like protein of bacteria in amino
C;Keywords: DNA binding; monomer

  Query Match             33.3%;  Score 11;  DB 2;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LEI 5
              :||
Db          1 MEI 3


RESULT 7
PT0644
T-cell receptor beta chain V-D-J region (111-1G) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0644
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0644
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             33.3%;  Score 11;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FT 2
              ||
Db          4 FT 5


RESULT 8
S19630
ribosomal protein L30 - Streptomyces griseus (fragment)
C;Species: Streptomyces griseus
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Jun-1997
C;Accession: S19630
R;Ochi, K.
Int. J. Syst. Bacteriol. 42, 144-150, 1992
A;Title: Electrophoretic heterogeneity of ribosomal protein AT-L30 among actinomycete ge
A;Reference number: S19630; MUID:92144363; PMID:1736962
A;Accession: S19630
A;Molecule type: protein
A;Residues: 1-7 <OCH>
A;Experimental source: strain IFO 13189
```

```
C;Superfamily: Escherichia coli ribosomal protein L30
C;Keywords: protein biosynthesis; ribosome

  Query Match             33.3%;  Score 11;  DB 2;  Length 7;
  Best Local Similarity   40.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LEISR 7
              |:|::
Db          3 LKITQ 7


RESULT 9
A60521
glycogen phosphorylase (EC 2.4.1.1), muscle - mullet (Liza ramada) (fragment)
N;Alternate names: glycogen phosphorylase b
C;Species: Liza ramada
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2003
C;Accession: A60521
R;Bonamusa, L.; Baanante, I.V.
Comp. Biochem. Physiol. B 95, 295-301, 1990
A;Title: Purification and characterization of glycogen phosphorylase B from skeletal mus
A;Reference number: A60521; MUID:90227907; PMID:2109669
A;Accession: A60521
A;Molecule type: protein
A;Residues: 1-5 <BON>
C;Superfamily: glucan phosphorylase
C;Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein
F;3/Binding site: phosphate (Ser) (covalent) (by phosphorylase b kinase) #status experim

  Query Match             30.3%;  Score 10;  DB 2;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EIS 6
              :||
Db          1 QIS 3


RESULT 10
B39127
phosphotransferase system enzyme II (EC 2.7.1.69) - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 08-Oct-1999
C;Accession: B39127
R;Hardesty, C.; Ferran, C.; DiRienzo, J.M.
J. Bacteriol. 173, 449-456, 1991
A;Title: Plasmid-mediated sucrose metabolism in Escherichia coli: characterization of sc
rin.
A;Reference number: A39127; MUID:91100329; PMID:1846143
A;Accession: B39127
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-7 <HAR>
A;Cross-references: GB:M38416; NID:g155142; PIDN:AAA98418.1; PID:g155144
C;Keywords: phosphotransferase

  Query Match             30.3%;  Score 10;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EIS 6
              :||
Db          5 QIS 7


RESULT 11
PS0254
18K protein 5507 - rice (strain Nihonbare) (fragment)
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Mar-1995
C;Accession: PS0254
R;Tsugita, A.
submitted to JIPID, April 1993
A;Reference number: PS0206
A;Accession: PS0254
A;Molecule type: protein
A;Residues: 1-7 <TSU>
A;Experimental source: leaf, chloroplast, strain Nihonbare
A;Note: molecular weight 18K, pI 4.4

  Query Match             30.3%;  Score 10;  DB 2;  Length 7;
  Best Local Similarity   40.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LEISR 7
              | |::
Db          1 LAIAK 5


RESULT 12
S33244
neuromodulatory peptide WWamide-1 - giant African snail
C;Species: Achatina fulica (giant African snail)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C;Accession: S33244
R;Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.
FEBS Lett. 323, 104-108, 1993
A;Title: WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of t
A;Reference number: S33244; MUID:93265912; PMID:8495720
A;Accession: S33244
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIN>

  Query Match             30.3%;  Score 10;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EIS 6
              |:|
Db          3 EMS 5


RESULT 13
S33245
neuromodulatory peptide WWamide-2 - giant African snail
C;Species: Achatina fulica (giant African snail)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C;Accession: S33245
R;Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.
FEBS Lett. 323, 104-108, 1993
A;Title: WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of th
A;Reference number: S33244; MUID:93265912; PMID:8495720
A;Accession: S33245
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIN>

  Query Match             30.3%;  Score 10;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EIS 6
              |:|
Db          3 EMS 5


RESULT 14
T13892
cytochrome-c oxidase (EC 1.9.3.1) chain I [imported] - river lamprey mitochondrion (fragm
C;Species: mitochondrion Lampetra fluviatilis (river lamprey)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: T13892
R;Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
```

;Title: The main features of the craniate mitochondrial DNA between the ND1 and the COI
;Reference number: Z17775; MUID:97398704; PMID:9254918
;Accession: T13892
;Status: preliminary; translated from GB/EMBL/DDBJ
;Molecule type: DNA
;Residues: 1-3 <DEL>
;Cross-references: EMBL:Y09528; NID:g2340016; PIDN:CAA70721.1; PID:g4379123
;Genetics:
;Genome: mitochondrion
;Note: COI
;Keywords: mitochondrion; oxidoreductase

Query Match 27.3%; Score 9; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TL 3
||
Db 2 TL 3

RESULT 15
E42364
flagellar protein fliR - Salmonella typhimurium (fragment)
C;Species: Salmonella typhimurium
C;Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 30-Sep-1993
C;Accession: E42364
R;Vogler, A.P.; Homma, M.; Irikura, V.M.; Macnab, R.M.
J. Bacteriol. 173, 3564-3572, 1991
A;Title: Salmonella typhimurium mutants defective in flagellar filament regrowth and seq
A;Reference number: A42364; MUID:91258342; PMID:1646201
A;Accession: E42364
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-5 <VOG>
A;Cross-references: GB:M62408

Query Match 27.3%; Score 9; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TL 3
||
Db 3 TL 4

Search completed: June 2, 2004, 19:00:23
Job time : 14 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:47:58 ; Search time 10 Seconds
(without alignments)
36.449 Million cell updates/sec

Title: US-09-712-819D-12
Perfect score: 33
Sequence: 1 FTLEISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 88

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	13	39.4	6	1	UN06_CLOPA	P81351 clostridium
2	11	33.3	5	1	RE31_LITRU	P82072 litoria rub
3	11	33.3	5	1	RE32_LITRU	P82073 litoria rub
4	10	30.3	7	1	EI05_LITRU	P82101 litoria rub
5	10	30.3	7	1	WWA1_ACHFU	P35919 achatina fu
6	10	30.3	7	1	WWA3_ACHFU	P35921 achatina fu
7	9	27.3	5	1	UF01_MOUSE	P38639 mus musculu
8	9	27.3	7	1	CCF1_ENTFA	P20104 enterococcu
9	9	27.3	7	1	CHOX_ALCSP	P16101 alcaligenes
10	9	27.3	7	1	CIA_ENTFA	P11932 enterococcu
11	9	27.3	7	1	GFRP_MOUSE	P99025 mus musculu
12	9	27.3	7	1	LANC_CARUI	P36960 carnobacter
13	8	24.2	5	1	AL14_CARMA	P81817 carcinus ma
14	8	24.2	5	1	PSK_DAUCA	P58261 daucus caro
15	8	24.2	5	1	UC22_MAIZE	P80628 zea mays (m
16	8	24.2	7	1	ALL2_CARMA	P81805 carcinus ma
17	8	24.2	7	1	ALL3_CARMA	P81806 carcinus ma
18	8	24.2	7	1	ALL4_CARMA	P81807 carcinus ma
19	8	24.2	7	1	ALL5_CARMA	P81808 carcinus ma
20	8	24.2	7	1	ALL7_CYDPO	P82158 cydia pomon
21	8	24.2	7	1	FAR1_ASCSU	P31889 ascaris suu
22	7	21.2	5	1	RE21_LITRU	P82071 litoria rub
23	7	21.2	6	1	ACPH_RABIT	P25154 oryctolagus
24	7	21.2	6	1	LOK1_LOCMI	P41491 locusta mig
25	7	21.2	7	1	FAR2_ASCSU	P31890 ascaris suu
26	7	21.2	7	1	UF03_MOUSE	P38641 mus musculu
27	7	21.2	7	1	WWA2_ACHFU	P35920 achatina fu
28	6	18.2	4	1	ACH1_ACHFU	P35904 achatina fu
29	6	18.2	4	1	FAR3_HIRME	P42562 hirudo medi
30	6	18.2	4	1	FAR4_HIRME	P42563 hirudo medi
31	6	18.2	4	1	FFKA_ANTEL	P58705 anthopleura
32	6	18.2	4	1	FLRF_HIRME	P42561 hirudo medi
33	6	18.2	4	1	FLRN_ANTEL	P58707 anthopleura
34	6	18.2	4	1	FMRF_MACNI	P01162 macrocallis
35	6	18.2	4	1	FYRI_ANTEL	P58706 anthopleura
36	6	18.2	4	1	OCP1_OCTMI	P58648 octopus min
37	6	18.2	5	1	BIOA_CITFR	P13071 citrobacter
38	6	18.2	5	1	EI03_LITRU	P82099 litoria rub
39	6	18.2	5	1	EI04_LITRU	P82100 litoria rub
40	6	18.2	5	1	FARP_ARTTR	P41853 artioposthi
41	6	18.2	5	1	PAP2_PARMA	P81864 pardachirus
42	6	18.2	5	1	RE11_LITRU	P82070 litoria rub
43	6	18.2	5	1	SUGA_ACHDO	P19991 acheta dome
44	6	18.2	5	1	TPIS_CANFA	P54714 canis famil
45	6	18.2	6	1	CIP1_MYTED	P13736 mytilus edu

ALIGNMENTS

RESULT 1
UN06_CLOPA
ID UN06_CLOPA STANDARD; PRT; 6 AA.
AC P81351;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Unknown protein CP 6 from 2D-page (Fragment).
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE.
RC STRAIN=W5;
RX MEDLINE=98291870; PubMed=9629918;
RA Flengsrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.0, ITS MW IS: 75.9 kDa.
FT NON_TER 6 6
SQ SEQUENCE 6 AA; 657 MW; 605B1DC1A45A8000 CRC64;

Query Match 39.4%; Score 13; DB 1; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 TLEI 5
              |||
Db          3 TAEI 6
```

RESULT 2
RE31_LITRU
ID RE31_LITRU STANDARD; PRT; 5 AA.
AC P82072;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rubellidin 3.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC Pelodryadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Skin secretion;
RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australin red tree frog
RT 'Litoria rubella'. The skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).

```
CC   -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC   -!- MASS SPECTROMETRY: MW=655; METHOD=FAB.
KW   Amphibian defense peptide; Amidation.
FT   MOD_RES       5      5       AMIDATION.
SQ   SEQUENCE   5 AA;  656 MW;  71A9C9CB10300000 CRC64;

  Query Match             33.3%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FT 2
              ||
Db          4 FT 5


RESULT 3
RE32_LITRU
ID   RE32_LITRU     STANDARD;      PRT;     5 AA.
AC   P82073;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Rubellidin 3.2.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Skin secretion;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litori electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:639-645(1999).
CC   -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
CC       activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
KW   Amphibian defense peptide.
SQ   SEQUENCE   5 AA;  570 MW;  71A9C9C862A00000 CRC64;

  Query Match             33.3%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FT 2
              ||
Db          4 FT 5


RESULT 4
EI05_LITRU
ID   EI05_LITRU     STANDARD;      PRT;     7 AA.
AC   P82101;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Electrin 5.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Skin secretion;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litori electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:639-645(1999).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Skin.
KW   Amphibian defense peptide; Amidation.
FT   MOD_RES       7      7       AMIDATION.
SQ   SEQUENCE   7 AA;  834 MW;  6DD05B076B0B5030 CRC64;

  Query Match             30.3%;  Score 10;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EIS 6
              ||:
Db          5 EIA 7


RESULT 5
WWA1_ACHFU
ID   WWA1_ACHFU     STANDARD;      PRT;     7 AA.
AC   P35919;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   WWamide-1.
OS   Achatina fulica (Giant African snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC   Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX   NCBI_TaxID=6530;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Ganglion;
RX   MEDLINE=93265912; PubMed=8495720;
RA   Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT   "WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT   ganglia of the African giant snail, Achatina fulica.";
RL   FEBS Lett. 323:104-108(1993).
CC   -!- FUNCTION: Exhibits modulatory effects on the peripheral nervous
CC       system. Inhibits activity on a central neuron.
DR   PIR; S33245; S33245.
KW   Neuropeptide; Amidation.
FT   MOD_RES       7      7       AMIDATION.
SQ   SEQUENCE   7 AA;  993 MW;  7362D5B69B041310 CRC64;

  Query Match             30.3%;  Score 10;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EIS 6
              |:|
Db          3 EMS 5


RESULT 6
WWA3_ACHFU
ID   WWA3_ACHFU     STANDARD;      PRT;     7 AA.
AC   P35921;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   WWamide-3.
OS   Achatina fulica (Giant African snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC   Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX   NCBI_TaxID=6530;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Ganglion;
RX   MEDLINE=93265912; PubMed=8495720;
RA   Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT   "WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT   ganglia of the African giant snail, Achatina fulica.";
```

```
RL   FEBS Lett. 323:104-108(1993).
DR   PIR; S33244; S33244.
KW   Neuropeptide; Amidation.
FT   MOD_RES       7      7       AMIDATION.
SQ   SEQUENCE   7 AA;  965 MW;  7362D5B69B132310 CRC64;

  Query Match             30.3%;  Score 10;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EIS 6
              |:|
Db          3 EMS 5


RESULT 7
UF01_MOUSE
ID   UF01_MOUSE     STANDARD;      PRT;     5 AA.
AC   P38639;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Unknown protein from 2D-page of fibroblasts (P19) (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Fibroblast;
RX   MEDLINE=95009907; PubMed=7523108;
RA   Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT   "Separation and sequencing of familiar and novel murine proteins
RT   using preparative two-dimensional gel electrophoresis.";
RL   Electrophoresis 15:735-745(1994).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 6.6, its MW is: 19 kDa.
FT   NON_TER       5      5
SQ   SEQUENCE   5 AA;  717 MW;  7364087043100000 CRC64;

  Query Match             27.3%;  Score 9;  DB 1;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ISR 7
              | |
Db          2 IGR 4


RESULT 8
CCF1_ENTFA
ID   CCF1_ENTFA     STANDARD;      PRT;     7 AA.
AC   P20104;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cCF10.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=89008313; PubMed=3139658;
RA   Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,
RA   Adsit J.C., Dunny G.M., Suzuki A.;
RT   "Structure of cCF10, a peptide sex pheromone which induces
RT   conjugative transfer of the Streptococcus faecalis tetracycline
RT   resistance plasmid, pCF10.";
RL   J. Biol. Chem. 263:14574-14578(1988).
CC   -!- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       HEMOLYSIN PLASMID PCF10.
DR   PIR; A30812; A30812.
KW   Pheromone.
SQ   SEQUENCE   7 AA;  790 MW;  72C9D2C731B2C740 CRC64;

  Query Match             27.3%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TL 3
              ||
Db          3 TL 4


RESULT 9
CHOX_ALCSP
ID   CHOX_ALCSP     STANDARD;      PRT;     7 AA.
AC   P16101;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-APR-1990 (Rel. 14, Last annotation update)
DE   Choline oxidase (EC 1.1.3.17) (Fragment).
OS   Alcaligenes sp.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Alcaligenes.
OX   NCBI_TaxID=512;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=81006769; PubMed=6997283;
RA   Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;
RT   "Identification and properties of the prosthetic group of choline
RT   oxidase from Alcaligenes sp.";
RL   J. Biochem. 88:197-203(1980).
CC   -!- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).
DR   PIR; A15398; A15398.
KW   Oxidoreductase.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  839 MW;  7415B1E457644AC0 CRC64;

  Query Match             27.3%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 SR 7
              ||
Db          6 SR 7


RESULT 10
CIA_ENTFA
ID   CIA_ENTFA      STANDARD;      PRT;     7 AA.
AC   P11932;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone cAM373 (Clumping-inducing agent) (CIA).
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=87005252; PubMed=3093276;
RA   Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
RA   White B.A., An F.Y., Clewell D.B., Suzuki A.;
RT   "Isolation and structure of the Streptococcus faecalis sex pheromone,
RT   cAM373.";
RL   FEBS Lett. 206:69-72(1986).
CC   -!- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
CC       HARBORING PAM373.
CC   -!- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
CC       SPECIFICITY OF PHEROMONES TO PLASMIDS.
CC   -!- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
DR   PIR; A25269; A25269.
KW   Pheromone.
SQ   SEQUENCE   7 AA;  734 MW;  75BDD72059C05DB0 CRC64;
```

```
  Query Match             27.3%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          3 FIL 5


RESULT 11
GFRP_MOUSE
ID   GFRP_MOUSE     STANDARD;      PRT;     7 AA.
AC   P99025;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
GN   GCHFR OR GFRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Liver;
RA   Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
RA   Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA   Cowthorne M.;
RL   Submitted (AUG-1998) to Swiss-Prot.
CC   -!- FUNCTION: Mediates tetrahydrobiopterin inhibition of GTP
CC       cyclohydrolase I. This inhibition is reversed by L-phenylalanine
CC       (By similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
DR   SWISS-2DPAGE; P99025; MOUSE.
FT   INIT_MET       0       0
FT   NON_TER        7       7
SQ   SEQUENCE   7 AA;  806 MW;  71B5B057273B4700 CRC64;

  Query Match             27.3%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 1.4e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LEIS 6
              | ||
Db          3 LLIS 6


RESULT 12
LANC_CARUI
ID   LANC_CARUI     STANDARD;      PRT;     7 AA.
AC   P36960;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lantibiotic carnocin UI49 (Fragment).
OS   Carnobacterium sp. (strain UI49).
OC   Bacteria; Firmicutes; Lactobacillales; Carnobacteriaceae;
OC   Carnobacterium.
OX   NCBI_TaxID=35782;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92321768; PubMed=1622206;
RA   Stoffels G., Nissen-Meyer J., Gudmundsdottir A., Sletten K., Holo H.,
RA   Nes I.F.;
RT   "Purification and characterization of a new bacteriocin isolated from
RT   a Carnobacterium sp.";
RL   Appl. Environ. Microbiol. 58:1417-1422(1992).
CC   -!- FUNCTION: Lanthionine-containing peptide antibiotic (lantibiotic).
CC       Active on Gram-positive bacteria.
KW   Antibiotic; Bacteriocin; Lantibiotic.
FT   NON_TER        7       7
SQ   SEQUENCE   7 AA;  786 MW;  741776D05B05B810 CRC64;

  Query Match             27.3%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EI 5
              ||
Db          3 EI 4


RESULT 13
AL14_CARMA
ID   AL14_CARMA     STANDARD;      PRT;     5 AA.
AC   P81817;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 14.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES        5       5       AMIDATION (POTENTIAL).
SQ   SEQUENCE   5 AA;  586 MW;  672879D5AB300000 CRC64;

  Query Match             24.2%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          3 FGL 5


RESULT 14
PSK_DAUCA
ID   PSK_DAUCA      STANDARD;      PRT;     5 AA.
AC   P58261;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Phytosulfokine-alpha (PSK-alpha) [Contains: Phytosulfokine-beta (PSK-
DE   beta)].
OS   Daucus carota (Carrot).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   campanulids; Apiales; Apiaceae; Apioideae; Scandiceae; Daucinae;
OC   Daucus.
OX   NCBI_TaxID=4039;
RN   [1]
RP   SEQUENCE, AND IDENTIFICATION BY MASS SPECTROMETRY.
RC   STRAIN=cv. US-Harumakigosun;
RX   MEDLINE=20212743; PubMed=10750705;
RA   Hanai H., Matsuno T., Yamamoto M., Matsubayashi Y., Kobayashi T.,
RA   Kamada H., Sakagami Y.;
RT   "A secreted peptide growth factor, phytosulfokine, acting as a
RT   stimulatory factor of carrot somatic embryo formation.";
RL   Plant Cell Physiol. 41:27-32(2000).
CC   -!- FUNCTION: In presence of 2,4-D, stimulates proliferation of the
CC       cells, but does not stimulate differentiation into the somatic
CC       embryos.
```

```
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: Sulfation is important for activity and for the binding to a
CC       putative membrane receptor (By similarity).
CC   -!- SIMILARITY: Belongs to the phytosulfokine family.
KW   Growth factor; Sulfation.
FT   PEPTIDE       1       4       PHYTOSULFOKINE-BETA.
FT   MOD_RES       1       1       SULFATION.
FT   MOD_RES       3       3       SULFATION.
SQ   SEQUENCE   5 AA;  687 MW;  76C1BB504B300000 CRC64;

  Query Match             24.2%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FT 2
              :|
Db          3 YT 4


RESULT 15
UC22_MAIZE
ID   UC22_MAIZE     STANDARD;      PRT;     5 AA.
AC   P80628;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Unknown protein from 2D-page of etiolated coleoptile (Spot 474)
DE   (Fragment).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Coleoptile;
RA   Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA   Pernollet J.-C., Zivy M., de Vienne D.;
RT   "The maize two dimensional gel protein database: towards an integrated
RT   genome analysis program.";
RL   Theor. Appl. Genet. 93:997-1005(1996).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 6.1, its MW is: 30.4 kDa.
DR   Maize-2DPAGE; P80628; COLEOPTILE.
DR   MaizeDB; 123954; -.
FT   NON_TER       1       1
FT   NON_TER       5       5
SQ   SEQUENCE   5 AA;  654 MW;  72CB19C9C0300000 CRC64;

  Query Match             24.2%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EI 5
              |:
Db          4 EV 5


Search completed: June  2, 2004, 18:58:22
Job time : 11 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:48:29 ; Search time 34.5 Seconds
(without alignments)
64.018 Million cell updates/sec

Title: US-09-712-819D-12
Perfect score: 33
Sequence: 1 FTLEISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 74

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	14	42.4	7	2	P83530	P83530 lactobacill
2	13	39.4	6	10	P82181	P82181 spinacia ol
3	13	39.4	6	10	P82182	P82182 spinacia ol
4	12	36.4	7	5	Q7Z1C0	Q7z1c0 caenorhabdi
5	9	27.3	7	2	O07354	O07354 synechococc
6	9	27.3	7	10	P93233	P93233 lycopersico
7	9	27.3	7	15	Q8JE81	Q8je81 human immun
8	8	24.2	7	2	P70804	P70804 azotobacter
9	8	24.2	7	4	Q15903	Q15903 homo sapien
10	8	24.2	7	12	Q9YVE3	Q9yve3 human adeno
11	8	24.2	7	12	Q9YIQ9	Q9yiq9 human adeno
12	8	24.2	7	12	Q9YQ10	Q9yq10 transmissib
13	8	24.2	7	12	Q9YIR0	Q9yir0 human adeno
14	8	24.2	7	13	Q8JJ20	Q8jj20 gallus gall
15	7	21.2	4	5	P83568	P83568 sepia offic
16	7	21.2	6	10	P82541	P82541 spinacia ol
17	7	21.2	7	2	Q47477	Q47477 escherichia
18	7	21.2	7	2	Q47029	Q47029 enterobacte
19	7	21.2	7	2	P72081	P72081 nocardia la
20	7	21.2	7	2	Q54248	Q54248 streptomyce
21	7	21.2	7	2	Q8KMS9	Q8kms9 enterobacte
22	7	21.2	7	4	Q15897	Q15897 homo sapien
23	7	21.2	7	8	Q95945	Q95945 saccharomyc
24	6	18.2	5	13	P83308	P83308 gallus gall
25	6	18.2	6	2	P83533	P83533 lactobacill
26	6	18.2	7	2	O50556	O50556 actinobacil
27	6	18.2	7	2	O34028	O34028 sphingomona
28	6	18.2	7	8	O98866	O98866 spinacia ol
29	6	18.2	7	10	P82445	P82445 nicotiana t
30	6	18.2	7	10	Q9C5B3	Q9c5b3 arabidopsis
31	6	18.2	7	11	Q8K3H6	Q8k3h6 rattus norv
32	6	18.2	7	12	Q66113	Q66113 cherry leaf
33	6	18.2	7	13	O42564	O42564 fugu rubrip
34	5	15.2	4	11	Q08433	Q08433 rattus sp.
35	5	15.2	5	2	P83073	P83073 bacillus ce
36	5	15.2	7	2	Q8KMS3	Q8kms3 klebsiella
37	5	15.2	7	2	Q47505	Q47505 escherichia
38	5	15.2	7	2	Q8GL12	Q8gl12 borrelia bu
39	5	15.2	7	2	Q8GL04	Q8gl04 borrelia bu
40	5	15.2	7	2	Q8GL00	Q8gl00 borrelia bu
41	5	15.2	7	3	P83492	P83492 bionectria
42	5	15.2	7	4	Q8NHH7	Q8nhh7 homo sapien
43	5	15.2	7	6	Q28742	Q28742 oryctolagus
44	5	15.2	7	8	P92214	P92214 amblyopyrum
45	5	15.2	7	8	P92393	P92393 hordeum vul

ALIGNMENTS

RESULT 1
P83530
ID P83530 PRELIMINARY; PRT; 7 AA.
AC P83530;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Unknown protein from 2D-page (Fragment).
OS Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1625;
RN [1]
RP SEQUENCE.
RC STRAIN=DSM 20451;
RX PubMed=12112860;
RA Drews O., Weiss W., Reil G., Parlar H., Wait R., Goerg A.;
RT "High pressure effects step-wise altered protein expression in
RT Lactobacillus sanfranciscensis.";
RL Proteomics 2:765-774(2002).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS UNKNOWN
CC PROTEIN IS: 15 KDA.
FT NON_TER 1 1
FT NON_TER 7 7
SQ SEQUENCE 7 AA; 790 MW; 6AAB02CAA731B2A0 CRC64;

Query Match 42.4%; Score 14; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLEI 5
||::
Db 2 TLDV 5

RESULT 2
P82181
ID P82181 PRELIMINARY; PRT; 6 AA.
AC P82181;

```
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 50S ribosomal protein L10 beta (Fragment).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435798; PubMed=10874046;
RA   Yamaguchi K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the 50 S subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 275:28466-28482(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC   -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002363; Ribosomal_L10eub.
DR   PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   NON_TER       6       6
SQ   SEQUENCE   6 AA;  675 MW;  6321B415B05DB000 CRC64;

  Query Match             39.4%;  Score 13;  DB 10;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ISR 7
              |||
Db          2 ISR 4


RESULT 3
P82182
ID   P82182      PRELIMINARY;      PRT;     6 AA.
AC   P82182;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 50S ribosomal protein L10 gamma (Fragment).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435798; PubMed=10874046;
RA   Yamaguchi K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the 50 S subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 275:28466-28482(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC   -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002363; Ribosomal_L10eub.
DR   PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   NON_TER       6       6
SQ   SEQUENCE   6 AA;  675 MW;  6321B415B05DB000 CRC64;

  Query Match             39.4%;  Score 13;  DB 10;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ISR 7
              |||
Db          2 ISR 4


RESULT 4
Q7Z1C0
ID   Q7Z1C0      PRELIMINARY;      PRT;     7 AA.
AC   Q7Z1C0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein W01B11.6.
GN   W01B11.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Wilson R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Bradshaw H., Graves T., Blair T.;
RT   "The sequence of C. elegans cosmid W01B11.";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Wilson R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF043704; AAO38592.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   7 AA;  874 MW;  72D1A9DB5041A6F0 CRC64;

  Query Match             36.4%;  Score 12;  DB 5;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              ||:
Db          5 FTV 7


RESULT 5
O07354
ID   O07354      PRELIMINARY;      PRT;     7 AA.
AC   O07354;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NifK (Fragment).
GN   NIFK.
OS   Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothece PCC 8801).
OC   Bacteria; Cyanobacteria; Chroococcales; Cyanothece.
OX   NCBI_TaxID=41431;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RF-1;
RX   MEDLINE=99231861; PubMed=10217509;
RA   Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RT   "Organization and expression of nitrogen-fixation genes in the aerobic
RT   nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
RT   RF-1.";
RL   Microbiology 145:743-753(1999).
DR   EMBL; AF003700; AAC35193.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  849 MW;  7412C72AA9D5B030 CRC64;

  Query Match             27.3%;  Score 9;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTL 3
              | |
Db          3 FDL 5


RESULT 6
P93233
ID   P93233      PRELIMINARY;      PRT;     7 AA.
AC   P93233;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14)
DE   (Fragment).
GN   LE-ACS1B.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97351561; PubMed=9207843;
RA   Oetiker J.H., Olson D.C., Shiu O.Y., Yang S.F.;
RT   "Differential induction of seven 1-aminocyclopropane-1-carboxylate
RT   synthase genes by elicitor in suspension cultures of tomato
RT   (Lycopersicon esculentum).";
RL   Plant Mol. Biol. 34:275-286(1997).
DR   EMBL; U75692; AAC49682.1; -.
DR   GO; GO:0016847; F:1-aminocyclopropane-1-carboxylate synthase . . .; IEA.
DR   GO; GO:0016829; F:lyase activity; IEA.
KW   Lyase.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  828 MW;  71B412C7377415D0 CRC64;

  Query Match             27.3%;  Score 9;  DB 10;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 SR 7
              ||
Db          1 SR 2


RESULT 7
Q8JE81
ID   Q8JE81      PRELIMINARY;      PRT;     7 AA.
AC   Q8JE81;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Truncated pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=4874;
RX   MEDLINE=22056123; PubMed=12060770;
RA   Beerenwinkel N., Schmidt B., Walter H., Kaiser R., Lengauer T.,
RA   Hoffmann D., Korn K., Selbig J.;
RT   "Diversity and complexity of HIV-1 drug resistance: A bioinformatics
RT   approach to predicting phenotype from genotype.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:8271-8276(2002).
DR   EMBL; AF347267; AAK32344.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  885 MW;  76C37731A046C700 CRC64;

  Query Match             27.3%;  Score 9;  DB 15;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TL 3
              ||
Db          4 TL 5


RESULT 8
P70804
ID   P70804      PRELIMINARY;      PRT;     7 AA.
AC   P70804;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AlgT protein (Fragment).
GN   ALGT.
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E;
RX   MEDLINE=96427318; PubMed=8830682;
RA   Rehm B.H.A., Ertesvag H., Valla S.;
RT   "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is
RT   part of an alg gene cluster physically organized in a manner similar
RT   to that in Pseudomonas aeruginosa.";
RL   J. Bacteriol. 178:5884-5889(1996).
DR   EMBL; X87973; CAA61230.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  684 MW;  71B5A5A5A2D1AED0 CRC64;

  Query Match             24.2%;  Score 8;  DB 2;  Length 7;
  Best Local Similarity   40.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TLEIS 6
              |:  |
Db          2 TVSSS 6


RESULT 9
Q15903
ID   Q15903      PRELIMINARY;      PRT;     7 AA.
AC   Q15903;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   (Clone XP7E7B) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
```

```
   Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
   Caskey C.T.H.;
   "Isolation of chromosome-specific genes by reciprocal probing of
   arrayed cDNAs and cosmid libraries.";
   Hum. Mol. Genet. 0:0-0(1995).
   EMBL; L32082; AAA73893.1; -.
   NON_TER       1      1
   NON_TER       7      7
   SEQUENCE   7 AA;  849 MW;  6B040339CDD33DB0 CRC64;

  Query Match             24.2%;  Score 8;  DB 4;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLE 4
              |  |
Db          4 FKRE 7


RESULT 10
Q9YVE3
ID   Q9YVE3      PRELIMINARY;      PRT;     7 AA.
AC   Q9YVE3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PVI core protein (Fragment).
GN   PVI.
OS   Human adenovirus type 7.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=10519;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Gomen;
RX   MEDLINE=99175282; PubMed=10074533;
RA   Crawford-Miksza L.K., Nang R.N., Schnurr D.P.;
RT   "Strain variation in adenovirus serotypes 4 and 7a causing acute
RT   respiratory disease.";
RL   J. Clin. Microbiol. 37:1107-1112(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Gomen;
RA   Crawford-Miksza L.K.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF065065; AAD03662.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  980 MW;  7B5EA414140322A0 CRC64;

  Query Match             24.2%;  Score 8;  DB 12;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ISR 7
              : |
Db          1 VKR 3


RESULT 11
Q9YIQ9
ID   Q9YIQ9      PRELIMINARY;      PRT;     7 AA.
AC   Q9YIQ9;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   PVI core protein (Fragment).
GN   PVI.
OS   Human adenovirus type 4.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
RT   "Molecular surveillance of strain variation in adenoviruses causing
RT   acute respiratory disease, AV 4 and AV 7a.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF065064; AAD03659.1; -.
DR   EMBL; AF065062; AAD03653.1; -.
DR   EMBL; AF065063; AAD03656.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  980 MW;  7B5EA414140322A0 CRC64;

  Query Match             24.2%;  Score 8;  DB 12;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ISR 7
              : |
Db          1 VKR 3


RESULT 12
Q9YQ10
ID   Q9YQ10      PRELIMINARY;      PRT;     7 AA.
AC   Q9YQ10;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical fusion protein.
OS   Transmissible gastroenteritis virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus.
OX   NCBI_TaxID=11149;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99099045; PubMed=9882359;
RA   Izeta A., Smerdou C., Alonso S., Penzes Z., Mendez A., Plana-Duran J.,
RA   Enjuanes L.;
RT   "Replication and packaging of transmissible gastroenteritis
RT   coronavirus-derived synthetic minigenomes.";
RL   J. Virol. 73:1535-1545(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95159435; PubMed=7856095;
RA   Eleouet J., Rasschaert D., Lambert P., Levy L., Vende P., Laude H.;
RT   "Complete sequence (20 kilobases) of the polyprotein-encoding gene 1
RT   of transmissible gastroenteritie virus.";
RL   Virology 206:817-822(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88078100; PubMed=2825819;
RA   Rasschaert D., Gelfi J., Laude H.;
RT   "Enteric coronavirus TGEV: partial sequence of the genomic RNA its
RT   organization and expression.";
RL   Biochimie 69:591-600(1987).
DR   EMBL; AJ011482; CAA09625.1; -.
SQ   SEQUENCE   7 AA;  927 MW;  69D6D7273B5726F0 CRC64;

  Query Match             24.2%;  Score 8;  DB 12;  Length 7;
  Best Local Similarity   25.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLE 4
              : |:
Db          3 YLLQ 6


RESULT 13
Q9YIR0
ID   Q9YIR0      PRELIMINARY;      PRT;     7 AA.
AC   Q9YIR0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   PVI core protein (Fragment).
```

```
OS   Human adenovirus type 7a.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=85755;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KN T96-0620, S-1058, and CL 68578;
RA   Crawford-Miksza L.K., Nang R.N., Schnurr D.P.;
RT   "Molecular surveillance of strain variation in adenoviruses causing
RT   acute respiratory disease, AV 4 and AV 7a.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF065068; AAD03668.1; -.
DR   EMBL; AF065066; AAD03664.1; -.
DR   EMBL; AF065067; AAD03666.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   7 AA;  980 MW;  7B5EA4141140322A0 CRC64;

  Query Match             24.2%;  Score 8;  DB 12;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ISR 7
              : |
Db          1 VKR 3


RESULT 14
Q8JJ20
ID   Q8JJ20      PRELIMINARY;      PRT;     7 AA.
AC   Q8JJ20;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Extracellular fatty acid binding protein (Fragment).
GN   EXFABP.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RA   Wang Q., Li N., Li H.;
RT   "Cloning and sequencing of 3' UTR of EXFABP gene in chicken.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF487519; AAL96665.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   7 AA;  780 MW;  72CB1AB2D5BEBB70 CRC64;

  Query Match             24.2%;  Score 8;  DB 13;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EI 5
              |:
Db          6 EV 7


RESULT 15
P83568
ID   P83568      PRELIMINARY;      PRT;     4 AA.
AC   P83568;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Pheromone peptide ILME.
OS   Sepia officinalis (Common cuttlefish).
OC   Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC   Decapodiformes; Sepioidea; Sepiidae; Sepia.
OX   NCBI_TaxID=6610;
RN   [1]
RP   SEQUENCE, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND MASS
RP   SPECTROMETRY.
RC   TISSUE=Egg;
RX   PubMed=10944467;
RA   Zatylny C., Gagnon J., Boucaud-Camou E., Henry J.;
RT   "ILME: a waterborne pheromonal peptide released by the eggs of Sepia
RT   officinalis.";
RL   Biochem. Biophys. Res. Commun. 275:217-222(2000).
RN   [2]
RP   SEQUENCE.
RC   TISSUE=Egg;
RX   PubMed=12207899;
RA   Zatylny C., Marvin L., Gagnon J., Henry J.;
RT   "Fertilization in Sepia officinalis: the first mollusk sperm-
RT   attracting peptide.";
RL   Biochem. Biophys. Res. Commun. 296:1186-1193(2002).
CC   -!- FUNCTION: HAS MYOTROPIC ACTIVITY TARGETING THE GENITAL TRACT.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- TISSUE SPECIFICITY: FOLLICLE, FULLY GROWN OOCYTE AND EGG(EC2).
CC   -!- MASS SPECTROMETRY: MW=505.4; METHOD=MALDI.
DR   GO; GO:0005186; F:pheromone activity; IEA.
KW   Pheromone.
SQ   SEQUENCE   4 AA;  505 MW;  6B16972030000000 CRC64;

  Query Match             21.2%;  Score 7;  DB 5;  Length 4;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LE 4
              :|
Db          3 ME 4


Search completed: June  2, 2004, 18:59:45
Job time : 36.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:47:08 ; Search time 49 Seconds
(without alignments)
40.364 Million cell updates/sec

Title: US-09-712-819D-12
Perfect score: 33
Sequence: 1 FTLEISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 92273

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	72.7	7	2	AAR81848	Aar81848 Human afa
2	24	72.7	7	2	AAY41889	Aay41889 Rheumatoi
3	24	72.7	7	4	ABB55870	Abb55870 Vascular
4	24	72.7	7	4	ABB56283	Abb56283 Vascular
5	24	72.7	7	4	ABB55981	Abb55981 Vascular
6	24	72.7	7	4	AAU28602	Aau28602 DPI trypt
7	24	72.7	7	4	AAU24969	Aau24969 Schizophr
8	24	72.7	7	4	AAU26249	Aau26249 Depressio
9	24	72.7	7	4	AAU15313	Aau15313 Schizophr
10	24	72.7	7	4	ABB52190	Abb52190 Human API
11	24	72.7	7	4	ABB52355	Abb52355 Human API
12	24	72.7	7	5	ABG78901	Abg78901 Multiple
13	24	72.7	7	5	ABG78730	Abg78730 Multiple
14	24	72.7	7	6	ABP58010	Abp58010 Prostate
15	24	72.7	7	6	ABP57255	Abp57255 Breast ca
16	24	72.7	7	6	ABP57203	Abp57203 Breast ca
17	24	72.7	7	6	ABR59010	Abr59010 Alzheimer
18	24	72.7	7	6	ABR59042	Abr59042 Alzheimer
19	23	69.7	7	2	AAY40736	Aay40736 S4 deriva
20	23	69.7	7	3	AAB30074	Aab30074 Scaffold
21	22	66.7	7	2	AAY40738	Aay40738 S4 deriva
22	22	66.7	7	3	AAB30076	Aab30076 Scaffold
23	20	60.6	6	6	ABR45678	Abr45678 Staphyloc
24	20	60.6	6	6	ABR46070	Abr46070 Staphyloc
25	20	60.6	6	6	ABR46854	Abr46854 Staphyloc
26	20	60.6	6	6	ABR44950	Abr44950 Staphyloc
27	20	60.6	6	6	ABR46462	Abr46462 Staphyloc
28	20	60.6	6	6	ABR45342	Abr45342 Staphyloc
29	20	60.6	7	2	AAR07656	Aar07656 Ribonucle
30	20	60.6	7	2	AAY42013	Aay42013 Rheumatoi
31	20	60.6	7	2	AAW82668	Aaw82668 Cauliflow
32	20	60.6	7	6	ABU09135	Abu09135 Human int
33	19	57.6	6	6	ABR45118	Abr45118 Staphyloc
34	19	57.6	6	6	ABR45846	Abr45846 Staphyloc
35	19	57.6	6	6	ABR46238	Abr46238 Staphyloc
36	19	57.6	6	6	ABR46630	Abr46630 Staphyloc
37	19	57.6	6	6	ABR47022	Abr47022 Staphyloc
38	19	57.6	6	6	ABR45454	Abr45454 Staphyloc
39	19	57.6	7	2	AAR72775	Aar72775 Mammalian
40	19	57.6	7	2	AAY40737	Aay40737 S4 deriva
41	19	57.6	7	3	AAY83858	Aay83858 Ribonucle
42	19	57.6	7	3	AAB30075	Aab30075 Scaffold
43	19	57.6	7	4	AAB47357	Aab47357 Cell lysi
44	19	57.6	7	7	ADB79629	Adb79629 Parapoxvi
45	18	54.5	6	2	AAW84431	Aaw84431 HIV-1 nuc

ALIGNMENTS

RESULT 1
AAR81848
ID AAR81848 standard; peptide; 7 AA.
XX
AC AAR81848;
XX
DT 16-MAY-1996 (first entry)
XX
DE Human afamin tryptic fragment FX20.
XX
KW Human; afamin; serum protein family; albumin; alpha-foetoprotein; plasma;
KW vitamin D binding protein; homology; post-translational processing;
KW chromatography; Primer; PCR; amplification; probe; rheumatoid arthritis;
KW ischaemia-reperfusion injury; ARDS; cardiopulmonary bypass; sepsis;
KW toxic plasma substance; inflammation.
XX
OS Homo sapiens.
XX
PN WO9527059-A1.
XX
PD 12-OCT-1995.
XX
PF 31-MAR-1995; 95WO-US004075.
XX
PR 31-MAR-1994; 94US-00222619.
XX
PA (AMGE-) AMGEN INC.
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Lichenstein HS, Lyons DE, Wurfel MM, Wright SD;
XX
DR WPI; 1995-358634/46.
XX
PT Human afamin or a variant and poly:nucleotide(s) encoding it - a human
PT serum protein with activities in common with other members of this
PT family.
XX
PS Example 3; Page 45; 97pp; English.
XX
CC Peptides AAR81847-54 are tryptic peptide fragments from human afamin
CC (AAR81845) novel member of the human serum protein family. The fragments
CC were used to design primers and probes (AAT00786-98) for the cloning of
CC the afamin gene (AAT00785) from human liver cDNA. Afamin is thought to
CC have similar properties to human albumin, alpha-foetoprotein and vitamin
CC D binding protein due to homology with these proteins. The gene encodes a
CC mature protein of 66576 daltons without post-translational processing
CC (ca. 87000 daltons with post-translational processing). The protein was
CC isolated from human plasma by a conventional chromatographic methods. The

CC protein can be used to ameliorate ischaemia-reperfusion injury,
CC rheumatoid arthritis, ARDS, cardiopulmonary bypass, sepsis, toxic plasma
CC substances released after inflammation, etc
XX
SQ Sequence 7 AA;

Query Match 72.7%; Score 24; DB 2; Length 7;
Best Local Similarity 71.4%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FTLEISR 7
|| | ||
Db 1 FTFEYSR 7

RESULT 2
AAY41889
ID AAY41889 standard; peptide; 7 AA.
XX
AC AAY41889;
XX
DT 09-DEC-1999 (first entry)
XX
DE Rheumatoid arthritis diagnostic protein isoform peptide #40.
XX
KW Human; rheumatoid arthritis; RA; diagnosis; RPI; RADF; detection;
KW rheumatoid arthritis diagnostic feature; ERPI; synovial fluid;
KW rheumatoid arthritis diagnostic protein isoform; screening;
KW expression reference protein isoform; prognosis.
XX
OS Homo sapiens.
XX
PN WO9947925-A2.
XX
PD 23-SEP-1999.
XX
PF 15-MAR-1999; 99WO-GB000763.
XX
PR 13-MAR-1998; 98GB-00005477.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Parekh RB, Patel TP, Townsend RR;
XX
DR WPI; 1999-571871/48.
XX
PT Diagnosis of human rheumatoid arthritis by two-dimensional
PT electrophoresis.
XX
PS Disclosure; Page 18; 157pp; English.
XX
CC A method has been developed for the diagnosis of human rheumatoid
CC arthritis (RA) using two-dimensional electrophoresis to generate a two-
CC dimensional array of features. The method can be used for screening,
CC diagnosis and prognosis of RA in a subject or for monitoring the effect
CC of an anti-RA drug or therapy administered to a subject. The method
CC comprises: (a) analysing a sample of serum or plasma and optionally
CC synovial fluid by two-dimensional electrophoresis, to generate a two-
CC dimensional array of features; (b) identifying at least one chosen
CC feature whose relative abundance correlates with the presence or absence
CC of RA; and (c) comparing the abundance of each chosen feature in the
CC sample with the abundance of that chosen feature in serum or plasma from
CC one or more persons without RA, where the relative abundance of the
CC chosen feature or features in the sample indicates the presence or
CC absence of RA in the subject. The method can also be used in clinical
CC studies for testing drugs for therapy of RA, for purification of RA-
CC diagnostic protein isoforms (RPIs), and for production of antibodies to
CC RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify
CC compounds that promote or inhibit their activity, which are then used as
CC RA drugs. Nucleic acid encoding RADFs can be used in gene therapy
CC protocols. AAY41844 to AAY42100 represent RPI peptides, AAY42101 to
CC AAY42103 represent expression reference protein isoform peptides and
CC AAZ25066 to AAZ25068 represent degenerate probes for RPIs, which are all
CC used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 72.7%; Score 24; DB 2; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLEISR 7
:| |:||
Db 1 YTFELSR 7

RESULT 3
ABB55870
ID ABB55870 standard; peptide; 7 AA.
XX
AC ABB55870;
XX
DT 15-FEB-2002 (first entry)
XX
DE Vascular dementia-associated protein isoform (VPI) 70.
XX
KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW diagnosis; prognosis; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200169261-A2.
XX
PD 20-SEP-2001.
XX
PF 14-MAR-2001; 2001WO-GB001106.
XX
PR 15-MAR-2000; 2000GB-00006285.
PR 24-NOV-2000; 2000GB-00028734.
PR 28-NOV-2000; 2000US-00724391.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-557937/62.
XX
PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT determining stage of VD and monitoring the effect of VD therapy,
PT comprises analyzing body fluid by 2-dimensional electrophoresis for
PT features correlated with VD.
XX
PS Claim 6; Page 31; 151pp; English.
XX
CC The invention relates to screening, diagnosis or prognosis of Vascular
CC Dementia (VD) in a subject comprising analysing body fluid from the
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC features containing at least one chosen feature whose relative abundance
CC correlates with the presence, absence, stage or severity of VD or
CC predicts the onset or course of VD, especially detecting in a sample of
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC specification. Detecting VD-associated features and VPI is useful for the
CC screening, diagnosis or prognosis of VD, for determining the stage or
CC severity of VD, for identifying a subject at risk of VD or for monitoring
CC the effect of therapy administered to a subject having VD. Nucleic acids
CC encoding a VPI or inhibiting the function of a VPI are useful for the
CC treatment of VD and for gene therapy
XX
SQ Sequence 7 AA;

Query Match 72.7%; Score 24; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLEISR 7

```
              :| |:||
Db          1 YTFELSR 7


RESULT 4
ABB56283
ID   ABB56283 standard; peptide; 7 AA.
XX
AC   ABB56283;
XX
DT   15-FEB-2002  (first entry)
XX
DE   Vascular dementia-associated protein isoform (VPI) 483.
XX
KW   Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW   diagnosis; prognosis; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200169261-A2.
XX
PD   20-SEP-2001.
XX
PF   14-MAR-2001; 2001WO-GB001106.
XX
PR   15-MAR-2000; 2000GB-00006285.
PR   24-NOV-2000; 2000GB-00028734.
PR   28-NOV-2000; 2000US-00724391.
XX
PA   (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI   Herath HMAC,  Parekh RB,  Rohlff C;
XX
DR   WPI; 2001-557937/62.
XX
PT   Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT   determining stage of VD and monitoring the effect of VD therapy,
PT   comprises analyzing body fluid by 2-dimensional electrophoresis for
PT   features correlated with VD.
XX
PS   Claim 6; Page 40; 151pp; English.
XX
CC   The invention relates to screening, diagnosis or prognosis of Vascular
CC   Dementia (VD) in a subject comprising analysing body fluid from the
CC   subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC   features containing at least one chosen feature whose relative abundance
CC   correlates with the presence, absence, stage or severity of VD or
CC   predicts the onset or course of VD, especially detecting in a sample of
CC   cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC   protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC   specification. Detecting VD-associated features and VPI is useful for the
CC   screening, diagnosis or prognosis of VD, for determining the stage or
CC   severity of VD, for identifying a subject at risk of VD or for monitoring
CC   the effect of therapy administered to a subject having VD. Nucleic acids
CC   encoding a VPI or inhibiting the function of a VPI are useful for the
CC   treatment of VD and for gene therapy
XX
SQ   Sequence 7 AA;

  Query Match             72.7%;  Score 24;  DB 4;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7


RESULT 5
ABB55981
ID   ABB55981 standard; peptide; 7 AA.
XX
AC   ABB55981;
XX
DT   15-FEB-2002  (first entry)
XX
DE   Vascular dementia-associated protein isoform (VPI) 181.
XX
KW   Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW   diagnosis; prognosis; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200169261-A2.
XX
PD   20-SEP-2001.
XX
PF   14-MAR-2001; 2001WO-GB001106.
XX
PR   15-MAR-2000; 2000GB-00006285.
PR   24-NOV-2000; 2000GB-00028734.
PR   28-NOV-2000; 2000US-00724391.
XX
PA   (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI   Herath HMAC,  Parekh RB,  Rohlff C;
XX
DR   WPI; 2001-557937/62.
XX
PT   Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT   determining stage of VD and monitoring the effect of VD therapy,
PT   comprises analyzing body fluid by 2-dimensional electrophoresis for
PT   features correlated with VD.
XX
PS   Claim 6; Page 33; 151pp; English.
XX
CC   The invention relates to screening, diagnosis or prognosis of Vascular
CC   Dementia (VD) in a subject comprising analysing body fluid from the
CC   subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC   features containing at least one chosen feature whose relative abundance
CC   correlates with the presence, absence, stage or severity of VD or
CC   predicts the onset or course of VD, especially detecting in a sample of
CC   cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC   protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC   specification. Detecting VD-associated features and VPI is useful for the
CC   screening, diagnosis or prognosis of VD, for determining the stage or
CC   severity of VD, for identifying a subject at risk of VD or for monitoring
CC   the effect of therapy administered to a subject having VD. Nucleic acids
CC   encoding a VPI or inhibiting the function of a VPI are useful for the
CC   treatment of VD and for gene therapy
XX
SQ   Sequence 7 AA;

  Query Match             72.7%;  Score 24;  DB 4;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7


RESULT 6
AAU28602
ID   AAU28602 standard; peptide; 7 AA.
XX
AC   AAU28602;
XX
DT   03-JAN-2002  (first entry)
XX
DE   DPI tryptic digest peptide #199.
XX
KW   Human; depression associated protein isoform; tryptic digest peptide;
KW   DPI; cerebrospinal fluid; CSF; BAD; bipolar affective disorder;
KW   neuropsychiatric disorder; bipolar mood disorder; neuroleptic;
KW   maniac-depressive illness; schizoaffective disorder.
```

XX
OS Homo sapiens.
XX
PN WO200162787-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000786.
XX
PR 24-FEB-2000; 2000GB-00004412.
PR 08-DEC-2000; 2000GB-00030050.
PR 12-DEC-2000; 2000US-0254830P.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C, Terrett JA, Tyson KL;
XX
DR WPI; 2001-570626/64.
XX
PT Novel nucleic acid encoding a protein associated with bipolar affective
PT disorder, which is used for diagnosis, prophylaxis and therapy of
PT neuropsychiatric disorders, such as bipolar affective disorder.
XX
PS Disclosure; Page 34; 153pp; English.
XX
CC The present invention relates to the identification of depression
CC associated protein isoforms (DPIs), particularly the tryptic digest
CC peptides of these proteins. Some of the DPIs (AAU28404-AAU28625)
CC described are decreased in the cerebrospinal fluid (CSF) of BAD (bipolar
CC affective disorder) subjects, whilst other DPIs (AAU28626-AAU28887) are
CC increased in BAD subjects. Also described are peptide sequences
CC identified from DPI-45 and DPI-213 and the nucleic acid sequence they are
CC encoded by. The sequences of the invention are useful for clinical
CC screening, diagnosis, prognosis, therapy and prophylaxis of
CC neuropsychiatric disorders e.g. BAD (also known as bipolar mood disorder,
CC BP), maniac-depressive illnesses, attention deficit disorders,
CC schizoaffective disorders, and unipolar affective disorders. The present
CC sequence represents one of the DPI tryptic digest peptides of the present
CC invention
XX
SQ Sequence 7 AA;

Query Match 72.7%; Score 24; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLEISR 7
:| |:||
Db 1 YTFELSR 7

RESULT 7
AAU24969
ID AAU24969 standard; peptide; 7 AA.
XX
AC AAU24969;
XX
DT 18-DEC-2001 (first entry)
XX
DE Schizophrenia-Associated Protein Isoform (SPI) peptide #198.
XX
KW Schizophrenia-associated protein isoform; SPI; SPI-206; SPI-238; SPI-240;
KW neuroleptic; gene therapy; cerebrospinal fluid; serum; plasma.
XX
OS Homo sapiens.
XX
PN WO200162785-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000792.
XX
PR 24-FEB-2000; 2000GB-00004415.
PR 28-DEC-2000; 2000US-00750395.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C, Terrett JA, Tyson KL;
XX
DR WPI; 2001-570624/64.
XX
PT New schizophrenia associated protein isoforms and encoding nucleic acid
PT molecules, useful for treatment, diagnosis and prognosis of schizophrenia
PT and screening for potential drugs for treatment and new drug targets.
XX
PS Disclosure; Page 32; 148pp; English.
XX
CC The sequence represents a schizophrenia-associated protein isoform (SPI).
CC These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable
CC in cerebrospinal fluid, serum or plasma and are useful markers of
CC schizophrenia. The sequences can be used for treatment and diagnosis of
CC schizophrenia, screening, prognosis, monitoring the results of therapy,
CC identifying patients most likely to respond to a particular therapy and
CC identification of new targets for drug treatment. SPI DNA is useful as a
CC nucleic acid probe to detect the presence of nucleic acids or SPIs
XX
SQ Sequence 7 AA;

Query Match 72.7%; Score 24; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLEISR 7
:| |:||
Db 1 YTFELSR 7

RESULT 8
AAU26249
ID AAU26249 standard; peptide; 7 AA.
XX
AC AAU26249;
XX
DT 18-DEC-2001 (first entry)
XX
DE Depression-Associated Protein isoform DPI-208.
XX
KW Human; Bipolar Affective Disorder; BAD; Depression-Associated feature;
KW DF; Depression-Associated protein isoform; DPI; Cerebro-spinal fluid;
KW CSF; antidepressant; antimanic; nootropic; tranquiliser; neuroleptic;
KW attention deficient disorder; schizoaffective disorder;
KW unipolar affective disorder.
XX
OS Homo sapiens.
XX
PN WO200163294-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000791.
XX
PR 24-FEB-2000; 2000GB-00004412.
PR 08-DEC-2000; 2000GB-00030050.
PR 12-DEC-2000; 2000US-0254830P.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-582081/65.
XX
PT Preparation for diagnosing or treating bipolar affected disorder (BAD) or
PT unipolar depression, or for screening for modulators, comprises a BAD-
PT associated protein isoform.
XX
PS Claim 8; Page 34; 163pp; English.

XX
CC The invention relates to a preparation comprising an isolated Bipolar
CC Affected Disorder (BAD)-Associated Protein Isoform (DPIs). The DPI's are
CC used to screen, diagnose or prognose of BAD or unipolar depression,
CC determine the stage or severity of BAD or unipolar depression, identify a
CC subject at risk of developing BAD or unipolar depression, or monitor the
CC effect of therapy in a subject. They are also used to screen for or
CC identify agents that interact with a DPI. These agents, antibodies
CC against the DPIs, and nucleic acids encoding the DPIs are used to treat
CC or prevent BAD or unipolar depression. Diseases that can be treated are
CC attention deficient disorder, a schizoaffective disorder, a bipolar or a
CC unipolar affective disorder. The DPIs are used in proteomics. The
CC proteomic approach of using DPIs for screening, diagnosis or prognosis of
CC BAD or unipolar depression overcomes the problems of using gene
CC expression analysis, such as not being able to obtain central nervous
CC system (CNS) tissue from a living patient under normal circumstances. The
CC present sequence is a DIP decreased in the CSF (cerebro-spinal fluid) of
CC subjects having BAD
XX
SQ Sequence 7 AA;

Query Match 72.7%; Score 24; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLEISR 7
:| |:||
Db 1 YTFELSR 7

RESULT 9
AAU15313
ID AAU15313 standard; peptide; 7 AA.
XX
AC AAU15313;
XX
DT 24-OCT-2001 (first entry)
XX
DE Schizophrenia-associated isoform peptide #198.
XX
KW Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder;
KW neurological disorder; neuropathy.
XX
OS Homo sapiens.
XX
PN WO200163293-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000783.
XX
PR 24-FEB-2000; 2000GB-00004415.
PR 28-DEC-2000; 2000US-00750395.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-502868/55.
XX
PT Diagnosing and monitoring Schizophrenia by detecting the presence of
PT Schizophrenia Associated Features and Schizophrenia Associated Protein
PT Isoforms in samples of cerebrospinal fluid.
XX
PS Claim 6; Page 32; 160pp; English.
XX
CC The invention relates to methods and compositions for screening,
CC diagnosis and prognosis of Schizophrenia. The method involves detecting
CC the presence of Schizophrenia (SCH) Associated Features (SFs) and SCH
CC Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis,
CC immunoassay or hybridisation assay, for diagnosing and monitoring SCH,
CC studying the effectiveness of treatments and for identifying potential
CC therapeutic agents. The method is used for (1) screening or diagnosis of
CC SCH and the relative abundance of at least 1 chosen feature correlates
CC with the presence or absence of SCH; and (2) monitoring the effect of
CC therapy administered to a subject with SCH and the relative abundance of
CC at least 1 chosen feature which correlates with the severity of SCH. The
CC expression and activity of the SFs, SPIs and related molecules (e.g.
CC secondary messengers) are studied to diagnose SCH, monitor the progress
CC of the disorder and the effectiveness of treatment and as targets to
CC identify and produce potential therapeutic agents for the treatment of
CC SCH. The paucity of detectable neuralgic defects distinguishes
CC neuropsychiatric disorders such as SCH from neurological disorders, where
CC manifestations of anatomical and biochemical changes have been identified
CC in many cases. Consequently the identification and characterisation of
CC cellular and/or molecular causative defects and neuropathies are
CC necessary for improved treatment of neuropsychiatric disorders. AAU15114-
CC AAU15762 represent the amino acid sequences of schizophrenia-associated
CC isoforms used in the method of the invention
XX
SQ Sequence 7 AA;

Query Match 72.7%; Score 24; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLEISR 7
:| |:||
Db 1 YTFELSR 7

RESULT 10
ABB52190
ID ABB52190 standard; peptide; 7 AA.
XX
AC ABB52190;
XX
DT 08-FEB-2002 (first entry)
XX
DE Human API-146 tryptic digest peptide #1.
XX
KW Human; neuroprotective; nootropic; gene therapy; vaccine;
KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;
KW Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;
KW Expression Reference Protein Isoform; ERPI; proteolysis.
XX
OS Homo sapiens.
XX
PN WO200175454-A2.
XX
PD 11-OCT-2001.
XX
PF 03-APR-2001; 2001WO-US010908.
XX
PR 03-APR-2000; 2000US-0194504P.
PR 28-NOV-2000; 2000US-0253647P.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
PA (PFIZ) PFIZER INC.
XX
PI Durham KL, Friedman DL, Herath HMAC, Kimmel LH, Parekh RB;
PI Potter DM, Rohlff C, Silber BM, Stiger TR, Sunderland PT;
PI Townsend RR, White F, Williams SA;
XX
DR WPI; 2001-639384/73.
XX
PT Screening for Alzheimer's disease in a mammal, by making two-dimensional
PT array of a feature whose relative abundance correlates with disease, and
PT comparing with abundance of the feature in samples of healthy persons.
XX
PS Example; Page 30; 162pp; English.
XX
CC The invention relates to methods for the screening, diagnosis and
CC prognosis of Alzheimer's disease. The methods involve the detection of
CC Alzheimer's Disease-Associated Features (AFs) and Alzheimer's Disease-
CC Associated Protein Isoforms (APIs) in cerebrospinal fluid, serum or

```
CC   plasma. The abundance of the AFs and APIs is then normalised to an
CC   Expression Reference Protein Isoform (ERPI) in order to determine whether
CC   a patient is suffering from, or has a predisposition to, Alzheimer's
CC   Disease. The relative abundance of the AFs and APIs correlates with the
CC   severity of Alzheimer's Disease. The present sequence is a peptide
CC   produced from an API by proteolysis
XX
SQ   Sequence 7 AA;

  Query Match             72.7%;  Score 24;  DB 4;  Length 7;
  Best Local Similarity   71.4%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              || | ||
Db          1 FTFEYSR 7


RESULT 11
ABB52355
ID   ABB52355 standard; peptide; 7 AA.
XX
AC   ABB52355;
XX
DT   08-FEB-2002  (first entry)
XX
DE   Human API-125 tryptic digest peptide #8.
XX
KW   Human; neuroprotective; nootropic; gene therapy; vaccine;
KW   Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;
KW   Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;
KW   Expression Reference Protein Isoform; ERPI; proteolysis.
XX
OS   Homo sapiens.
XX
PN   WO200175454-A2.
XX
PD   11-OCT-2001.
XX
PF   03-APR-2001; 2001WO-US010908.
XX
PR   03-APR-2000; 2000US-0194504P.
PR   28-NOV-2000; 2000US-0253647P.
XX
PA   (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
PA   (PFIZ ) PFIZER INC.
XX
PI   Durham KL,  Friedman DL,  Herath HMAC,  Kimmel LH,  Parekh RB;
PI   Potter DM,  Rohlff C,  Silber BM,  Stiger TR,  Sunderland PT;
PI   Townsend RR,  White F,  Williams SA;
XX
DR   WPI; 2001-639384/73.
XX
PT   Screening for Alzheimer's disease in a mammal, by making two-dimensional
PT   array of a feature whose relative abundance correlates with disease, and
PT   comparing with abundance of the feature in samples of healthy persons.
XX
PS   Example; Page 34; 162pp; English.
XX
CC   The invention relates to methods for the screening, diagnosis and
CC   prognosis of Alzheimer's disease. The methods involve the detection of
CC   Alzheimer's Disease-Associated Features (AFs) and Alzheimer's Disease-
CC   Associated Protein Isoforms (APIs) in cerebrospinal fluid, serum or
CC   plasma. The abundance of the AFs and APIs is then normalised to an
CC   Expression Reference Protein Isoform (ERPI) in order to determine whether
CC   a patient is suffering from, or has a predisposition to, Alzheimer's
CC   Disease. The relative abundance of the AFs and APIs correlates with the
CC   severity of Alzheimer's Disease. The present sequence is a peptide
CC   produced from an API by proteolysis
XX
SQ   Sequence 7 AA;

  Query Match             72.7%;  Score 24;  DB 4;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7


RESULT 12
ABG78901
ID   ABG78901 standard; peptide; 7 AA.
XX
AC   ABG78901;
XX
DT   29-NOV-2002  (first entry)
XX
DE   Multiple sclerosis associated feature (MSF) tryptic digest peptide #389.
XX
KW   Multiple sclerosis; MS; multiple sclerosis associated feature; MSAF;
KW   human; multiple sclerosis-associated protein isoform; MSPI;
KW   antiinflammatory; neuroprotective.
XX
OS   Homo sapiens.
XX
PN   WO200259604-A2.
XX
PD   01-AUG-2002.
XX
PF   25-JAN-2002; 2002WO-GB000330.
XX
PR   26-JAN-2001; 2001US-0264404P.
PR   20-NOV-2001; 2001US-0331647P.
XX
PA   (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI   Herath HMAC,  Parekh RB,  Rohlff C;
XX
DR   WPI; 2002-599812/64.
XX
PT   Screening or diagnosing multiple sclerosis (MS), useful for e.g.
PT   determining the stage or severity of MS, comprises detecting the presence
PT   of MS-associated features or protein isoforms by 2-dimensional
PT   electrophoresis.
XX
PS   Disclosure; Page 39; 128pp; English.
XX
CC   This invention relates to a novel method for screening or diagnosing
CC   multiple sclerosis (MS) in a subject to determine the stage or severity
CC   of MS, to identify a subject at risk of developing MS or to monitor the
CC   effect of a therapy administered. The method comprises analysing a sample
CC   body fluid from the subject by two-dimensional electrophoresis and
CC   detecting the presence of multiple sclerosis-associated features (MSFs),
CC   or multiple sclerosis-associated protein isoforms (MSPIs). The MSF's of
CC   the invention correspond to spots identified on a 2D gel these proteins
CC   may have antiinflammatory or neuroprotective activity. The methods of the
CC   invention and the compositions are useful for clinical screening,
CC   diagnosis and treatment of MS, for monitoring the effectiveness of MS
CC   treatment, for selecting participants in clinical trials, for identifying
CC   patients most likely to respond to a particular therapeutic treatment and
CC   for screening and developing drugs for treatment of MS. Agents that
CC   modulate the expression or activity of an MSPI are useful for treating
CC   MS, for preventing or delaying the onset or development of MS, to prevent
CC   or delay the progression of MS, or to ameliorate the symptoms MS. Nucleic
CC   acids comprising a sequence encoding an MSPI, MSPI-related polypeptide,
CC   or their fragments are useful for promoting MSPI function by gene
CC   therapy. The present sequence represents a human multiple sclerosis
CC   associated feature tryptic digest peptide of the invention
XX
SQ   Sequence 7 AA;

  Query Match             72.7%;  Score 24;  DB 5;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7
```

RESULT 13
ABG78730
ID ABG78730 standard; peptide; 7 AA.
XX
AC ABG78730;
XX
DT 29-NOV-2002 (first entry)
XX
DE Multiple sclerosis associated feature (MSF) tryptic digest peptide #218.
XX
KW Multiple sclerosis; MS; multiple sclerosis associated feature; MSAF;
KW human; multiple sclerosis-associated protein isoform; MSPI;
KW antiinflammatory; neuroprotective.
XX
OS Homo sapiens.
XX
PN WO200259604-A2.
XX
PD 01-AUG-2002.
XX
PF 25-JAN-2002; 2002WO-GB000330.
XX
PR 26-JAN-2001; 2001US-0264404P.
PR 20-NOV-2001; 2001US-0331647P.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2002-599812/64.
XX
PT Screening or diagnosing multiple sclerosis (MS), useful for e.g.
PT determining the stage or severity of MS, comprises detecting the presence
PT of MS-associated features or protein isoforms by 2-dimensional
PT electrophoresis.
XX
PS Disclosure; Page 26; 128pp; English.
XX
CC This invention relates to a novel method for screening or diagnosing
CC multiple sclerosis (MS) in a subject to determine the stage or severity
CC of MS, to identify a subject at risk of developing MS or to monitor the
CC effect of a therapy administered. The method comprises analysing a sample
CC body fluid from the subject by two-dimensional electrophoresis and
CC detecting the presence of multiple sclerosis-associated features (MSFs),
CC or multiple sclerosis-associated protein isoforms (MSPIs). The MSF's of
CC the invention correspond to spots identified on a 2D gel these proteins
CC may have antiinflammatory or neuroprotective activity. The methods of the
CC invention and the compositions are useful for clinical screening,
CC diagnosis and treatment of MS, for monitoring the effectiveness of MS
CC treatment, for selecting participants in clinical trials, for identifying
CC patients most likely to respond to a particular therapeutic treatment and
CC for screening and developing drugs for treatment of MS. Agents that
CC modulate the expression or activity of an MSPI are useful for treating
CC MS, for preventing or delaying the onset or development of MS, to prevent
CC or delay the progression of MS, or to ameliorate the symptoms MS. Nucleic
CC acids comprising a sequence encoding an MSPI, MSPI-related polypeptide,
CC or their fragments are useful for promoting MSPI function by gene
CC therapy. The present sequence represents a human multiple sclerosis
CC associated feature tryptic digest peptide of the invention
XX
SQ Sequence 7 AA;

```
  Query Match             72.7%;  Score 24;  DB 5;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7
```

RESULT 14
ABP58010
ID ABP58010 standard; peptide; 7 AA.
XX
AC ABP58010;
XX
DT 11-FEB-2003 (first entry)
XX
DE Prostate cancer marker protein peptide fragment.
XX
KW Prostate cancer; marker; vitamin D binding protein; VDBP; human;
KW diagnosis; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200275314-A2.
XX
PD 26-SEP-2002.
XX
PF 30-NOV-2001; 2001WO-US045031.
XX
PR 30-NOV-2000; 2000US-0250284P.
PR 08-NOV-2001; 2001US-0344948P.
XX
PA (MATR-) MATRITECH INC.
XX
PI Hlavaty J, Briggman JV;
XX
DR WPI; 2003-067369/06.
XX
PT Diagnosing or treating prostrate cancer by detecting in a sample isolated
PT from the individual the presence of prostate cancer-associated protein.
XX
PS Claim 1; Page 41; 63pp; English.
XX
CC The present sequence is that of a peptide fragment of a novel human 50.8
CC kDa prostate cancer-associated protein that has been identified as a
CC highly effective marker for prostate cancer. The novel protein includes a
CC polypeptide that is related to human serum vitamin D binding protein
CC (VDBP, see ABP58017). The present peptide corresponds to amino acids 346-
CC 352 of this VDBP allele. It is one of a series of peptides (see ABP58005-
CC 16) that distinguish VDBP-related proteins from other proteins, or which
CC may be characterised as binding specifically to an anti-VDBP antibody.
CC VDBP-related proteins are detectable at a higher concentration in serum
CC from a mammal, e.g. a human, with prostate cancer relative to serum from
CC a healthy mammal and can therefore be used as prostate cancer markers.
CC They permit the rapid detection, preferably before metastases occur, of
CC prostate cancer. A target prostate cancer-associated protein may be
CC detected using a labelled antibody capable of binding specifically to the
CC protein. Prostate cancer-associated proteins, and nucleic acids encoding
CC them, are also useful as targets for treating prostate cancer, and as
CC indicators for monitoring the efficiency of prostate cancer therapy
XX
SQ Sequence 7 AA;

```
  Query Match             72.7%;  Score 24;  DB 6;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7
```

RESULT 15
ABP57255
ID ABP57255 standard; peptide; 7 AA.
XX
AC ABP57255;

XX
DT 16-APR-2003 (first entry)
XX
DE Breast cancer associated tryptic digest peptide SEQ ID NO:154.
XX
KW Breast cancer associated feature; BF; BPI; breast cancer; diagnosis;
KW breast cancer associated protein isoform; cytostatic; gene therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200288750-A2.
XX
PD 07-NOV-2002.
XX
PF 02-MAY-2002; 2002WO-GB002022.
XX
PR 02-MAY-2001; 2001GB-00010790.
PR 27-JUL-2001; 2001GB-00018385.
PR 14-AUG-2001; 2001GB-00019791.
PR 16-AUG-2001; 2001GB-00020045.
PR 22-NOV-2001; 2001GB-00028062.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC;
XX
DR WPI; 2003-175048/17.
XX
PT Screening, diagnosing or determining the stage or severity of breast
PT cancer, comprises analyzing and quantitatively detecting Breast Cancer-
PT Associated Features or Breast Cancer-Associated Protein Isoforms in a
PT biological sample.
XX
PS Example; Page 77; 88pp; English.
XX
CC The present invention describes a method for screening, diagnosing or
CC determining the stage or severity of breast cancer, identifying a subject
CC at risk of developing breast cancer, or monitoring the effect of therapy
CC administered to a subject with breast cancer, by generating a two-
CC dimensional array of features comprising breast cancer-associated
CC features (BFs), or quantitatively detecting breast cancer-associated
CC protein isoforms (BPIs). Also described: (1) an antibody capable of
CC immunospecifically binding to one of the BPIs; (2) a pharmaceutical
CC compositions comprising: (a) a BPI, or a nucleic acid encoding a BPI, and
CC a carrier; or (b) the antibody of (1), or a fragment or derivative of the
CC antibody, and a carrier; (3) screening for agents that interact with one
CC or more BPIs, BPI fragments, polypeptides related to BPIs, or BPI-fusion
CC proteins; (4) screening for or identifying agents that modulate the
CC expression or activity of one or more BPIs, a BPI fragment, a BPI-related
CC polypeptide, or BPI-fusion proteins; and (5) treating or preventing
CC breast cancer. BPIs have cytostatic activity and can be used in gene
CC therapy. Methods and kits comprising antibodies or the BPIs from the
CC present invention can be used for screening, diagnosing or determining
CC the stage or severity of breast cancer, identifying a subject at risk of
CC developing breast cancer, or monitoring the effect of therapy
CC administered to a subject with breast cancer. The antibodies, BPIs,
CC nucleic acids encoding the BPIs, or an agent that modulates the activity
CC of one or more BPIs are useful for treating or preventing breast cancer.
CC ABP57104 to ABP57250 represent breast cancer associated tryptic digest
CC peptides, which are used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 72.7%; Score 24; DB 6; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7
```

Search completed: June 2, 2004, 18:57:50
Job time : 51 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:59:52 ; Search time 37.5 Seconds
(without alignments)
52.517 Million cell updates/sec

Title: US-09-712-819D-12
Perfect score: 33
Sequence: 1 FTLEISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 49349

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	72.7	7	9	US-09-791-378-198	Sequence 198, App
2	24	72.7	7	9	US-09-998-909-7	Sequence 7, Appli
3	24	72.7	7	9	US-09-826-290-205	Sequence 205, App
4	24	72.7	7	9	US-09-826-290-370	Sequence 370, App
5	24	72.7	7	10	US-09-791-393-145	Sequence 145, App
6	24	72.7	7	10	US-09-791-389-145	Sequence 145, App
7	24	72.7	7	12	US-10-601-100-105	Sequence 105, App
8	24	72.7	7	16	US-10-264-309-131	Sequence 131, App
9	24	72.7	7	16	US-10-264-309-460	Sequence 460, App
10	20	60.6	7	13	US-10-050-552A-4	Sequence 4, Appli
11	20	60.6	7	14	US-10-234-026-6	Sequence 6, Appli
12	20	60.6	7	15	US-10-238-965-4	Sequence 4, Appli
13	18	54.5	6	9	US-09-876-388-6	Sequence 6, Appli
14	18	54.5	6	14	US-10-105-930-39	Sequence 39, Appl
15	18	54.5	6	14	US-10-287-892-6	Sequence 6, Appli
16	18	54.5	6	14	US-10-288-340-6	Sequence 6, Appli
17	18	54.5	7	9	US-09-772-105-23	Sequence 23, Appl
18	18	54.5	7	9	US-09-876-388-7	Sequence 7, Appli
19	18	54.5	7	14	US-10-287-892-7	Sequence 7, Appli
20	18	54.5	7	14	US-10-288-340-7	Sequence 7, Appli
21	17	51.5	6	9	US-09-727-963A-38	Sequence 38, Appl
22	17	51.5	6	12	US-10-328-953-42	Sequence 42, Appl
23	17	51.5	6	12	US-10-367-580-312	Sequence 312, App
24	17	51.5	6	12	US-10-367-593-312	Sequence 312, App
25	17	51.5	6	12	US-10-367-594-312	Sequence 312, App
26	17	51.5	6	12	US-10-367-654-312	Sequence 312, App
27	17	51.5	6	12	US-10-367-658-312	Sequence 312, App
28	17	51.5	6	12	US-10-367-668-312	Sequence 312, App
29	17	51.5	6	14	US-10-006-869-636	Sequence 636, App
30	17	51.5	6	14	US-10-006-869-650	Sequence 650, App
31	17	51.5	6	14	US-10-020-354-101	Sequence 101, App
32	17	51.5	6	15	US-10-395-032-636	Sequence 636, App
33	17	51.5	6	15	US-10-395-032-650	Sequence 650, App
34	17	51.5	7	9	US-09-832-312-77	Sequence 77, Appl
35	17	51.5	7	9	US-09-734-417-12	Sequence 12, Appl
36	17	51.5	7	11	US-09-829-495-77	Sequence 77, Appl
37	17	51.5	7	12	US-10-396-678-5	Sequence 5, Appli
38	17	51.5	7	14	US-10-006-869-637	Sequence 637, App
39	17	51.5	7	14	US-10-006-869-639	Sequence 639, App
40	17	51.5	7	14	US-10-006-869-651	Sequence 651, App
41	17	51.5	7	14	US-10-006-869-653	Sequence 653, App
42	17	51.5	7	15	US-10-395-032-637	Sequence 637, App
43	17	51.5	7	15	US-10-395-032-639	Sequence 639, App
44	17	51.5	7	15	US-10-395-032-651	Sequence 651, App
45	17	51.5	7	15	US-10-395-032-653	Sequence 653, App

ALIGNMENTS

RESULT 1
US-09-791-378-198
; Sequence 198, Application US/09791378
; Patent No. US20020142303A1
; GENERAL INFORMATION:
; APPLICANT: Parekh, Rajesh
; TITLE OF INVENTION: PROTEINS, GENES AND THEIR USE FOR DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: SCHIZOPHRENIA
; FILE REFERENCE: 9195-061-999
; CURRENT APPLICATION NUMBER: US/09/791,378
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/750,395
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 677
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 198
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-378-198

Query Match 72.7%; Score 24; DB 9; Length 7;
Best Local Similarity 57.1%; Pred. No. 1e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7
```

RESULT 2
US-09-998-909-7
; Sequence 7, Application US/09998909
; Patent No. US20020164664A1
; GENERAL INFORMATION:
; APPLICANT: Hlavaty, John
; APPLICANT: Briggman, Joseph
; TITLE OF INVENTION: Detection and Treatment of Prostate Cancer

```
;  FILE REFERENCE: MTP-027
;  CURRENT APPLICATION NUMBER: US/09/998,909
;  CURRENT FILING DATE:  2001-11-30
;  PRIOR APPLICATION NUMBER: US 60/250,284
;  PRIOR FILING DATE: 2000-11-30
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-998-909-7

  Query Match           72.7%;  Score 24;  DB 9;  Length 7;
  Best Local Similarity  57.1%;  Pred. No. 1e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7


RESULT 3
US-09-826-290-205
; Sequence 205, Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
;  APPLICANT: Durham, L.Kathryn
;  APPLICANT:  Friedman, David L.
;  APPLICANT:  Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT:  Kimmel, Lida H.
;  APPLICANT:  Parekh, Rajesh Bhikhu
;  APPLICANT:  Potter, David M.
;  APPLICANT:  Rohlff, Christian
;  APPLICANT:  Silber, B. Michael
;  APPLICANT:  Stiger, Thomas R.
;  APPLICANT:  Sunderland, P. Trey
;  APPLICANT:  Townsend, Robert Reid
;  APPLICANT:  White, Frost
;  APPLICANT:  Williams, Stephen A.
;  TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
;  TITLE OF INVENTION:  Uses Therefor, Including Diagnosis and Treatment of
;  TITLE OF INVENTION:  Alzheimer's Disease
;  FILE REFERENCE: 2572-1-001 N2
;  CURRENT APPLICATION NUMBER: US/09/826,290
;  CURRENT FILING DATE:  2001-04-30
;  PRIOR APPLICATION NUMBER: US 60/194,504
;  PRIOR FILING DATE: 2000-04-03
;  PRIOR APPLICATION NUMBER: US 60/253,647
;  PRIOR FILING DATE: 2000-11-28
;  NUMBER OF SEQ ID NOS: 492
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 205
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-826-290-205

  Query Match           72.7%;  Score 24;  DB 9;  Length 7;
  Best Local Similarity  71.4%;  Pred. No. 1e+06;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              || | ||
Db          1 FTFEYSR 7


RESULT 4
US-09-826-290-370
; Sequence 370, Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
;  APPLICANT: Durham, L.Kathryn
;  APPLICANT:  Friedman, David L.
;  APPLICANT:  Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT:  Kimmel, Lida H.
;  APPLICANT:  Parekh, Rajesh Bhikhu
;  APPLICANT:  Potter, David M.
;  APPLICANT:  Rohlff, Christian
;  APPLICANT:  Silber, B. Michael
;  APPLICANT:  Stiger, Thomas R.
;  APPLICANT:  Sunderland, P. Trey
;  APPLICANT:  Townsend, Robert Reid
;  APPLICANT:  White, Frost
;  APPLICANT:  Williams, Stephen A.
;  TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
;  TITLE OF INVENTION:  Uses Therefor, Including Diagnosis and Treatment of
;  TITLE OF INVENTION:  Alzheimer's Disease
;  FILE REFERENCE: 2572-1-001 N2
;  CURRENT APPLICATION NUMBER: US/09/826,290
;  CURRENT FILING DATE:  2001-04-30
;  PRIOR APPLICATION NUMBER: US 60/194,504
;  PRIOR FILING DATE: 2000-04-03
;  PRIOR APPLICATION NUMBER: US 60/253,647
;  PRIOR FILING DATE: 2000-11-28
;  NUMBER OF SEQ ID NOS: 492
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 370
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-826-290-370

  Query Match           72.7%;  Score 24;  DB 9;  Length 7;
  Best Local Similarity  57.1%;  Pred. No. 1e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7


RESULT 5
US-09-791-393-145
; Sequence 145, Application US/09791393
; Publication No. US20030032200A1
; GENERAL INFORMATION:
;  APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT: Parekh, Rajesh Bhikhu
;  APPLICANT: Rohlff, Christian
;  TITLE OF INVENTION: Proteins, Genes and Their Use for
;  TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
;  TITLE OF INVENTION: and Unipolar Depression
;  FILE REFERENCE: 2543-1-001 N1
;  CURRENT APPLICATION NUMBER: US/09/791,393
;  CURRENT FILING DATE: 2002-01-02
;  EARLIER APPLICATION NUMBER: GB 0004412.3
;  EARLIER FILING DATE: 2000-02-24
;  EARLIER APPLICATION NUMBER: GB 0030050.9
;  EARLIER FILING DATE: 2000-12-08
;  EARLIER APPLICATION NUMBER: US 60/254,830
;  EARLIER FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 145
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-791-393-145

  Query Match           72.7%;  Score 24;  DB 10;  Length 7;
  Best Local Similarity  57.1%;  Pred. No. 1e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              :| |:||
```

```
Db          1 YTFELSR 7


RESULT 6
US-09-791-389-145
; Sequence 145, Application US/09791389
; Publication No. US20030032773A1
; GENERAL INFORMATION:
;  APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT:  Parekh, Rajesh Bhikhu
;  APPLICANT:  Rohlff, Christian
;  APPLICANT:  Terrett, Jonathan Alexander
;  APPLICANT:  Tyson, Kerry Louise
;  TITLE OF INVENTION: Proteins, Genes and Their Use for
;  TITLE OF INVENTION:  Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
;  TITLE OF INVENTION:  and Unipolar Depression
;  FILE REFERENCE: 2543-1-001 N2
;  CURRENT APPLICATION NUMBER: US/09/791,389
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: GB 0004412.3
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: GB 0030050.9
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: US 60/254,830
;  PRIOR FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 145
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-791-389-145

  Query Match             72.7%;  Score 24;  DB 10;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7


RESULT 7
US-10-601-100-105
; Sequence 105, Application US/10601100
; Publication No. US20040072261A1
; GENERAL INFORMATION:
;  APPLICANT: INNOGENETICS N.V.
;  TITLE OF INVENTION: Method for the Diagnosis and Differential Diagnosis of
;  TITLE OF INVENTION:  Neurological Diseases
;  FILE REFERENCE: 11362.0038.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/601,100
;  CURRENT FILING DATE:  2003-06-20
;  PRIOR APPLICATION NUMBER: EP 02447121.1
;  PRIOR FILING DATE: 2002-06-21
;  PRIOR APPLICATION NUMBER: US 60/396,437
;  PRIOR FILING DATE: 2002-07-17
;  NUMBER OF SEQ ID NOS: 113
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 105
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-10-601-100-105

  Query Match             72.7%;  Score 24;  DB 12;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7


RESULT 8
US-10-264-309-131
; Sequence 131, Application US/10264309
; Publication No. US20040022794A1
; GENERAL INFORMATION:
;  APPLICANT: DURHAM, L. KATHRYN
;  APPLICANT:  FRIEDMAN, DAVID L.
;  APPLICANT:  HERATH, HERATH
;  APPLICANT:  KIMMEL, LIDA H.
;  APPLICANT:  PAREKH, RAJESH B.
;  APPLICANT:  POTTER, DAVID M.
;  APPLICANT:  ROHLFF, CHRISTIAN
;  APPLICANT:  SILBER, B. MICHAEL
;  APPLICANT:  SNYDER, PETER J.
;  APPLICANT:  SOARES, HOLLY D.
;  APPLICANT:  STIGER, THOMAS R.
;  APPLICANT:  SUNDERLAND, P. TREY
;  APPLICANT:  TOWNSEND, ROBERT R.
;  APPLICANT:  WHITE, W. FROST
;  APPLICANT:  WILLIAMS, STEPHEN A.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES, POLYPEPTIDES AND USES THEREFOR,
;  TITLE OF INVENTION:  INCLUDING DIAGNOSIS AND TREATMENT OF ALZHEIMER'S DISEASE
;  FILE REFERENCE: POA-002.01
;  CURRENT APPLICATION NUMBER: US/10/264,309
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: 60/326,708
;  PRIOR FILING DATE: 2001-10-03
;  NUMBER OF SEQ ID NOS: 491
;  SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 131
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-264-309-131

  Query Match             72.7%;  Score 24;  DB 16;  Length 7;
  Best Local Similarity   71.4%;  Pred. No. 1e+06;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              || | ||
Db          1 FTFEYSR 7


RESULT 9
US-10-264-309-460
; Sequence 460, Application US/10264309
; Publication No. US20040022794A1
; GENERAL INFORMATION:
;  APPLICANT: DURHAM, L. KATHRYN
;  APPLICANT:  FRIEDMAN, DAVID L.
;  APPLICANT:  HERATH, HERATH
;  APPLICANT:  KIMMEL, LIDA H.
;  APPLICANT:  PAREKH, RAJESH B.
;  APPLICANT:  POTTER, DAVID M.
;  APPLICANT:  ROHLFF, CHRISTIAN
;  APPLICANT:  SILBER, B. MICHAEL
;  APPLICANT:  SNYDER, PETER J.
;  APPLICANT:  SOARES, HOLLY D.
;  APPLICANT:  STIGER, THOMAS R.
;  APPLICANT:  SUNDERLAND, P. TREY
;  APPLICANT:  TOWNSEND, ROBERT R.
;  APPLICANT:  WHITE, W. FROST
;  APPLICANT:  WILLIAMS, STEPHEN A.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES, POLYPEPTIDES AND USES THEREFOR,
;  TITLE OF INVENTION:  INCLUDING DIAGNOSIS AND TREATMENT OF ALZHEIMER'S DISEASE
;  FILE REFERENCE: POA-002.01
;  CURRENT APPLICATION NUMBER: US/10/264,309
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: 60/326,708
;  PRIOR FILING DATE: 2001-10-03
;  NUMBER OF SEQ ID NOS: 491
```

```
;  SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 460
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-264-309-460

  Query Match           72.7%;  Score 24;  DB 16;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              :| |:||
Db          1 YTFELSR 7


RESULT 10
US-10-050-552A-4
; Sequence 4, Application US/10050552A
; Publication No. US20020187512A1
; GENERAL INFORMATION:
;  APPLICANT: Nagem, Ronaldo A.P.
;  APPLICANT:  Colau, Didier
;  APPLICANT:  Renauld, Jean-Christophe
;  APPLICANT:  Dumoutier, Laure
;  APPLICANT:  Polikarpov, Igor
;  TITLE OF INVENTION: Crystal Structure of Interleukin-22
;  FILE REFERENCE: LUD-5722 US
;  CURRENT APPLICATION NUMBER: US/10/050,552A
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 60/317,937
;  PRIOR FILING DATE: 2001-09-10
;  PRIOR APPLICATION NUMBER: 60/333,150
;  PRIOR FILING DATE: 2001-11-27
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: A conserved sequence between Region 2 in IL-22 and IL-10
US-10-050-552A-4

  Query Match           60.6%;  Score 20;  DB 13;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLE 4
              ||||
Db          1 FTLE 4


RESULT 11
US-10-234-026-6
; Sequence 6, Application US/10234026
; Publication No. US20030097679A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut voor Biotechnol
;  APPLICANT:  Guy Jerome Corneel Bauw
;  APPLICANT:  Mark William Davey
;  APPLICANT:  Jens Ostergaard
;  APPLICANT:  Marc Charles Ernest Van Montagu
;  TITLE OF INVENTION: PRODUCTION OF ASCORBIC ACID IN PLANTS
;  FILE REFERENCE: DECLE5.001C1
;  CURRENT APPLICATION NUMBER: US/10/234,026
;  CURRENT FILING DATE:  2000-08-29
;  PRIOR APPLICATION NUMBER: NL 1006000
;  PRIOR FILING DATE: 1997-05-07
;  PRIOR APPLICATION NUMBER: PCT/EP98/02830
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 09/423,468
;  PRIOR FILING DATE: 2000-02-15
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Brassica oleracea
US-10-234-026-6

  Query Match           60.6%;  Score 20;  DB 14;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLEISR 7
              |:|:|:
Db          2 TIELSK 7


RESULT 12
US-10-238-965-4
; Sequence 4, Application US/10238965
; Publication No. US20040002586A1
; GENERAL INFORMATION:
;  APPLICANT: Nagem, Ronaldo A.P.
;  APPLICANT:  Colau, Didier
;  APPLICANT:  Renauld, Jean-Christophe
;  APPLICANT:  Dumoutier, Laure
;  APPLICANT:  Polikarpov, Igor
;  TITLE OF INVENTION: Crystal Structure of Interleukin-22
;  FILE REFERENCE: LUD-5722 US
;  CURRENT APPLICATION NUMBER: US/10/238,965
;  CURRENT FILING DATE:  2002-09-10
;  PRIOR APPLICATION NUMBER: 60/317,937
;  PRIOR FILING DATE: 2001-09-10
;  PRIOR APPLICATION NUMBER: 60/333,150
;  PRIOR FILING DATE: 2001-11-27
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: A conserved sequence between Region 2 in IL-22 and IL-10
US-10-238-965-4

  Query Match           60.6%;  Score 20;  DB 15;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLE 4
              ||||
Db          1 FTLE 4


RESULT 13
US-09-876-388-6
; Sequence 6, Application US/09876388
; Patent No. US20020049153A1
; GENERAL INFORMATION:
;  APPLICANT: Bridon, Dominique P.
;  APPLICANT:  L'Archeveque, Benoit
;  APPLICANT:  Ezrin, Alan M.
;  APPLICANT:  Holmes, Darren L.
;  APPLICANT:  Leblanc, Anouk
;  APPLICANT:  St. Pierre, Serge
;  TITLE OF INVENTION: LONG LASTING INSULINOTROPIC PEPTIDES
;  FILE REFERENCE: 500862001610
;  CURRENT APPLICATION NUMBER: US/09/876,388
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 09/623,618
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: PCT/US00/13563
;  PRIOR FILING DATE: 2000-05-17
```

; PRIOR APPLICATION NUMBER: 60/159,783
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/134,406
; PRIOR FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-876-388-6

Query Match 54.5%; Score 18; DB 9; Length 6;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLEIS 6
|| ::|
Db 1 FTSDVS 6

RESULT 14
US-10-105-930-39
; Sequence 39, Application US/10105930
; Publication No. US20030009018A1
; GENERAL INFORMATION:
; APPLICANT: Maeda, Masatsugu
; APPLICANT: Yaguchi, No. US20030009018A1iko
; TITLE OF INVENTION: NOVEL HEMOPOIETIN RECEPTOR PROTEIN, NR12
; FILE REFERENCE: 06501-105US1
; CURRENT APPLICATION NUMBER: US/10/105,930
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: PCT/JP00/06654
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: JP 2000-240397
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 11-273358
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-105-930-39

Query Match 54.5%; Score 18; DB 14; Length 6;
Best Local Similarity 60.0%; Pred. No. 1e+06;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLEI 5
:||:|
Db 1 YTLQI 5

RESULT 15
US-10-287-892-6
; Sequence 6, Application US/10287892
; Publication No. US20030108567A1
; GENERAL INFORMATION:
; APPLICANT: Bridon, Dominique P.
; APPLICANT: L'Archeveque, Benoit
; APPLICANT: Ezrin, Alan M.
; APPLICANT: Holmes, Darren L.
; APPLICANT: Leblanc, Anouk
; APPLICANT: St. Pierre, Serge
; TITLE OF INVENTION: LONG LASTING SYNTHETIC GLUCAGON LIKE PEPTIDE (GLP-1)
; FILE REFERENCE: 500862001612
; CURRENT APPLICATION NUMBER: US/10/287,892
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 09/657,332
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/159,783
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-287-892-6

Query Match 54.5%; Score 18; DB 14; Length 6;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLEIS 6
|| ::|
Db 1 FTSDVS 6

Search completed: June 2, 2004, 19:08:54
Job time : 37.5 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:56:07 ; Search time 15.5 Seconds
(without alignments)
23.315 Million cell updates/sec

Title: US-09-712-819D-12
Perfect score: 33
Sequence: 1 FTLEISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 57228

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	72.7	7	1	US-08-222-619-8	Sequence 8, Appli
2	24	72.7	7	5	PCT-US95-04075-8	Sequence 8, Appli
3	20	60.6	7	4	US-09-423-468A-6	Sequence 6, Appli
4	19	57.6	6	1	US-08-222-619-21	Sequence 21, Appl
5	19	57.6	6	5	PCT-US95-04075-21	Sequence 21, Appl
6	19	57.6	7	1	US-08-136-743B-55	Sequence 55, Appl
7	19	57.6	7	3	US-09-040-216-28	Sequence 28, Appl
8	18	54.5	6	1	US-08-297-731-4	Sequence 4, Appli
9	18	54.5	6	4	US-09-623-618B-6	Sequence 6, Appli
10	18	54.5	6	4	US-09-657-332A-6	Sequence 6, Appli
11	18	54.5	6	4	US-09-876-388-6	Sequence 6, Appli
12	18	54.5	6	5	PCT-US95-10793-4	Sequence 4, Appli
13	18	54.5	7	1	US-08-297-731-5	Sequence 5, Appli
14	18	54.5	7	4	US-09-623-618B-7	Sequence 7, Appli
15	18	54.5	7	4	US-09-218-363-23	Sequence 23, Appl
16	18	54.5	7	4	US-09-657-332A-7	Sequence 7, Appli
17	18	54.5	7	4	US-09-876-388-7	Sequence 7, Appli
18	18	54.5	7	4	US-08-753-750B-20	Sequence 20, Appl
19	18	54.5	7	5	PCT-US95-10793-5	Sequence 5, Appli
20	17	51.5	5	1	US-08-136-743B-63	Sequence 63, Appl
21	17	51.5	5	3	US-09-040-216-55	Sequence 55, Appl
22	17	51.5	5	3	US-08-591-632-23	Sequence 23, Appl
23	17	51.5	5	4	US-09-611-451-23	Sequence 23, Appl
24	17	51.5	6	1	US-08-136-743B-62	Sequence 62, Appl
25	17	51.5	6	3	US-09-040-216-54	Sequence 54, Appl
26	17	51.5	6	4	US-09-187-859-636	Sequence 636, App
27	17	51.5	6	4	US-09-187-859-650	Sequence 650, App
28	17	51.5	6	4	US-09-522-433B-12	Sequence 12, Appl
29	17	51.5	6	4	US-09-839-542B-636	Sequence 636, App
30	17	51.5	6	4	US-09-839-542B-650	Sequence 650, App
31	17	51.5	7	1	US-08-136-743B-6	Sequence 6, Appli
32	17	51.5	7	1	US-08-136-743B-29	Sequence 29, Appl
33	17	51.5	7	1	US-08-136-743B-33	Sequence 33, Appl
34	17	51.5	7	1	US-08-136-743B-37	Sequence 37, Appl
35	17	51.5	7	1	US-08-136-743B-54	Sequence 54, Appl
36	17	51.5	7	1	US-08-136-743B-56	Sequence 56, Appl
37	17	51.5	7	1	US-08-136-743B-57	Sequence 57, Appl
38	17	51.5	7	1	US-08-136-743B-58	Sequence 58, Appl
39	17	51.5	7	1	US-08-136-743B-59	Sequence 59, Appl
40	17	51.5	7	1	US-08-136-743B-60	Sequence 60, Appl
41	17	51.5	7	1	US-08-175-471-2	Sequence 2, Appli
42	17	51.5	7	2	US-08-429-054A-2	Sequence 2, Appli
43	17	51.5	7	2	US-08-718-777-2	Sequence 2, Appli
44	17	51.5	7	3	US-09-040-216-3	Sequence 3, Appli
45	17	51.5	7	3	US-09-040-216-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-222-619-8
; Sequence 8, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-222-619-8

Query Match 72.7%; Score 24; DB 1; Length 7;
Best Local Similarity 71.4%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FTLEISR 7
|| | ||
Db 1 FTFEYSR 7

RESULT 2
PCT-US95-04075-8
; Sequence 8, Application PC/TUS9504075
; GENERAL INFORMATION:

```
;    APPLICANT:  AMGEN INC.
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04075
;      FILING DATE:
;      CLASSIFICATION:
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US95-04075-8

  Query Match             72.7%;  Score 24;  DB 5;  Length 7;
  Best Local Similarity   71.4%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLEISR 7
              || | ||
Db          1 FTFEYSR 7


RESULT 3
US-09-423-468A-6
; Sequence 6, Application US/09423468A
; Patent No. 6469149
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut voor Biotechnol
;  APPLICANT:  Guy Jerome Corneel Bauw
;  APPLICANT:  Mark William Davey
;  APPLICANT:  Jens Ostergaard
;  APPLICANT:  Marc Charles Ernest Van Montagu
;  TITLE OF INVENTION: PRODUCTION OF ASCORBIC ACID IN PLANTS
;  FILE REFERENCE: DECLE5.001APC
;  CURRENT APPLICATION NUMBER: US/09/423,468A
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: NL 1006000
;  PRIOR FILING DATE: 1997-05-07
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Brassica oleracea
US-09-423-468A-6

  Query Match             60.6%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLEISR 7
              |:|:|:
Db          2 TIELSK 7


RESULT 4
US-08-222-619-21
; Sequence 21, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
;    APPLICANT:  Lichenstein, Henri
;    APPLICANT:  Lyons, David
;    APPLICANT:  Wurfel, Mark
;    APPLICANT:  Wright, Samuel
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/222,619
;      FILING DATE:
;      CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-222-619-21

  Query Match             57.6%;  Score 19;  DB 1;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 3e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLEIS 6
              || | |
Db          1 FTFEYS 6


RESULT 5
PCT-US95-04075-21
; Sequence 21, Application PC/TUS9504075
;  GENERAL INFORMATION:
;    APPLICANT:  AMGEN INC.
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04075
;      FILING DATE:
;      CLASSIFICATION:
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
```

```
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US95-04075-21

  Query Match             57.6%;  Score 19;  DB 5;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 3e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLEIS 6
              || | |
Db          1 FTPEYS 6


RESULT 6
US-08-136-743B-55
; Sequence 55, Application US/08136743B
; Patent No. 5459063
;  GENERAL INFORMATION:
;    APPLICANT:  Barry S. Cooperman, Harvey Rubin,
;    APPLICANT:  Jerome Salem, and Alison L. Fisher
;    TITLE OF INVENTION:  "Plasmodium falciparum Ribonu-
;    TITLE OF INVENTION:  cleotide Reductase, DNA Sequences Therefor and Peptide inhibit
;    TITLE OF INVENTION:  Thereof"
;    NUMBER OF SEQUENCES:  67
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The University of Pennsylvania
;      STREET:  Suite 330
;      STREET:  3700 Market Street
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19104-3246
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/136,743B
;      FILING DATE:  10/14/93
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  3957-10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5459063e
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-136-743B-55

  Query Match             57.6%;  Score 19;  DB 1;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLEI 5
              |||::
Db          1 FTLDL 5


RESULT 7
US-09-040-216-28
; Sequence 28, Application US/09040216
; Patent No. 6030942
; GENERAL INFORMATION:
;  APPLICANT: COOPERMAN, ET AL., BARRY
;  TITLE OF INVENTION: PEPTIDES, PEPTIDE ANALOGS, PEPTIDOMIMETICS, AND OTHER
;  TITLE OF INVENTION: SMALL MOLECULES USEFUL FOR INHIBITING THE ACTIVITY OF
;  TITLE OF INVENTION: RIBONUCLEOTIDE REDUCTASE
;  FILE REFERENCE: 9596-63U1
;  CURRENT APPLICATION NUMBER: US/09/040,216
;  CURRENT FILING DATE: 1998-03-17
;  EARLIER APPLICATION NUMBER: 08/919,748
;  EARLIER FILING DATE: 1997-08-28
;  EARLIER APPLICATION NUMBER: 60/025,146
;  EARLIER FILING DATE: 1996-08-30
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: Ribonucleotide reductase inhibitor peptide
;   FEATURE:
;   OTHER INFORMATION: residue 1:  MOD_RES:  ACETYLATION
US-09-040-216-28

  Query Match             57.6%;  Score 19;  DB 3;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLEI 5
              |||::
Db          1 FTLDL 5


RESULT 8
US-08-297-731-4
; Sequence 4, Application US/08297731
; Patent No. 5574008
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, William T.
;    APPLICANT:  Yakubu-Madus, Fatima E.
;    TITLE OF INVENTION:  BIOLOGICALLY ACTIVE FRAGMENTS OF
;    TITLE OF INVENTION:  GLUCAGON-LIKE INSULINOTROPIC PEPTIDE
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company/RSM
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/297,731
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maciak, Ronald S.
;      REGISTRATION NUMBER:  35,262
;      REFERENCE/DOCKET NUMBER:  X9630
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-1664
;      TELEFAX:  317-277-1917
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-297-731-4

  Query Match             54.5%;  Score 18;  DB 1;  Length 6;
```

Best Local Similarity 50.0%; Pred. No. 3e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTLEIS 6
              || ::|
Db          1 FTSDVS 6
```

RESULT 9
US-09-623-618B-6
; Sequence 6, Application US/09623618B
; Patent No. 6329336
; GENERAL INFORMATION:
; APPLICANT: Bridon, Dominique P.
; APPLICANT: L'Archeveque, Benoit
; APPLICANT: Ezrin, Alan M.
; APPLICANT: Holmes, Darren L.
; APPLICANT: Leblanc, Anouk
; APPLICANT: St. Pierre, Serge
; TITLE OF INVENTION: LONG LASTING INSULINOTROPIC PEPTIDES
; FILE REFERENCE: 500862001620
; CURRENT APPLICATION NUMBER: US/09/623,618B
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: PCT/US00/13563
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: 60/159,783
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/134,406
; PRIOR FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-623-618B-6

Query Match 54.5%; Score 18; DB 4; Length 6;
Best Local Similarity 50.0%; Pred. No. 3e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTLEIS 6
              || ::|
Db          1 FTSDVS 6
```

RESULT 10
US-09-657-332A-6
; Sequence 6, Application US/09657332A
; Patent No. 6514500
; GENERAL INFORMATION:
; APPLICANT: Bridon, Dominique P.
; APPLICANT: L'Archeveque, Benoit
; APPLICANT: Ezrin, Alan M.
; APPLICANT: Holmes, Darren L.
; APPLICANT: Leblanc, Anouk
; APPLICANT: St. Pierre, Serge
; TITLE OF INVENTION: LONG LASTING SYNTHETIC GLUCAGON LIKE PEPTIDE (GLP-1)
; FILE REFERENCE: 500862001600
; CURRENT APPLICATION NUMBER: US/09/657,332A
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/159,783
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/134,406
; PRIOR FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-657-332A-6

Query Match 54.5%; Score 18; DB 4; Length 6;
Best Local Similarity 50.0%; Pred. No. 3e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTLEIS 6
              || ::|
Db          1 FTSDVS 6
```

RESULT 11
US-09-876-388-6
; Sequence 6, Application US/09876388
; Patent No. 6593295
; GENERAL INFORMATION:
; APPLICANT: Bridon, Dominique P.
; APPLICANT: L'Archeveque, Benoit
; APPLICANT: Ezrin, Alan M.
; APPLICANT: Holmes, Darren L.
; APPLICANT: Leblanc, Anouk
; APPLICANT: St. Pierre, Serge
; TITLE OF INVENTION: LONG LASTING INSULINOTROPIC PEPTIDES
; FILE REFERENCE: 500862001610
; CURRENT APPLICATION NUMBER: US/09/876,388
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/623,618
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: PCT/US00/13563
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: 60/159,783
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/134,406
; PRIOR FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-876-388-6

Query Match 54.5%; Score 18; DB 4; Length 6;
Best Local Similarity 50.0%; Pred. No. 3e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FTLEIS 6
              || ::|
Db          1 FTSDVS 6
```

RESULT 12
PCT-US95-10793-4
; Sequence 4, Application PC/TUS9510793
; GENERAL INFORMATION:
; APPLICANT: Johnson, William T.
; APPLICANT: Yakubu-Madus, Fatima E.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE FRAGMENTS OF
; TITLE OF INVENTION: GLUCAGON-LIKE INSULINOTROPIC PEPTIDE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company/RSM
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: USA

```
;       ZIP:  46285
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/10793
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maciak, Ronald S.
;       REGISTRATION NUMBER:  35,262
;       REFERENCE/DOCKET NUMBER:  X9630
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  317-276-1664
;       TELEFAX:  317-277-1917
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
PCT-US95-10793-4

  Query Match             54.5%;  Score 18;  DB 5;  Length 6;
  Best Local Similarity   50.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEIS 6
              || ::|
Db          1 FTSDVS 6


RESULT 13
US-08-297-731-5
; Sequence 5, Application US/08297731
; Patent No. 5574008
;  GENERAL INFORMATION:
;     APPLICANT:  Johnson, William T.
;     APPLICANT:  Yakubu-Madus, Fatima E.
;     TITLE OF INVENTION:  BIOLOGICALLY ACTIVE FRAGMENTS OF
;     TITLE OF INVENTION:  GLUCAGON-LIKE INSULINOTROPIC PEPTIDE
;     NUMBER OF SEQUENCES:  13
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Eli Lilly and Company/RSM
;       STREET:  Lilly Corporate Center
;       CITY:  Indianapolis
;       STATE:  IN
;       COUNTRY:  USA
;       ZIP:  46285
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/297,731
;       FILING DATE:
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maciak, Ronald S.
;       REGISTRATION NUMBER:  35,262
;       REFERENCE/DOCKET NUMBER:  X9630
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  317-276-1664
;       TELEFAX:  317-277-1917
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-297-731-5

  Query Match             54.5%;  Score 18;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEIS 6
              || ::|
Db          2 FTSDVS 7


RESULT 14
US-09-623-618B-7
; Sequence 7, Application US/09623618B
; Patent No. 6329336
; GENERAL INFORMATION:
;  APPLICANT: Bridon, Dominique P.
;  APPLICANT:  L'Archeveque, Benoit
;  APPLICANT:  Ezrin, Alan M.
;  APPLICANT:  Holmes, Darren L.
;  APPLICANT:  Leblanc, Anouk
;  APPLICANT:  St. Pierre, Serge
;  TITLE OF INVENTION: LONG LASTING INSULINOTROPIC PEPTIDES
;  FILE REFERENCE: 500862001620
;  CURRENT APPLICATION NUMBER: US/09/623,618B
;  CURRENT FILING DATE:  2000-09-05
;  PRIOR APPLICATION NUMBER: PCT/US00/13563
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/159,783
;  PRIOR FILING DATE: 1999-10-15
;  PRIOR APPLICATION NUMBER: 60/134,406
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-09-623-618B-7

  Query Match             54.5%;  Score 18;  DB 4;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLEIS 6
              || ::|
Db          2 FTSDVS 7


RESULT 15
US-09-218-363-23
; Sequence 23, Application US/09218363
; Patent No. 6387616
; GENERAL INFORMATION:
;  APPLICANT: Ozelius, Laurie J.
;  APPLICANT: Breakefield, Xandra O.
;  TITLE OF INVENTION: TORSIN, TORSIN GENES, AND METHODS OF USE
;  FILE REFERENCE: MGH-1184pA2
;  CURRENT APPLICATION NUMBER: US/09/218,363
;  CURRENT FILING DATE: 1998-12-22
;  EARLIER APPLICATION NUMBER: 09/099,454
;  EARLIER FILING DATE: 1998-06-18
;  EARLIER APPLICATION NUMBER: 60/050,244
;  EARLIER FILING DATE: 1997-06-19
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
;   LENGTH: 7
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-218-363-23

  Query Match             54.5%;  Score 18;  DB 4;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLE 4
              ||:|
Db          2 FTME 5


Search completed: June  2, 2004, 19:01:06
Job time : 15.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:54:42 ; Search time 13 Seconds
(without alignments)
51.795 Million cell updates/sec

Title: US-09-712-819D-13
Perfect score: 31
Sequence: 1 LTLKLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 457

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	13	41.9	7	2	S19630	ribosomal protein
2	11	35.5	3	3	T13892	cytochrome-c oxida
3	11	35.5	4	2	T46627	hypothetical prote
4	11	35.5	5	2	E42364	flagellar protein
5	11	35.5	6	2	I49424	cytotoxic T-lympho
6	11	35.5	7	2	JN0859	peptidyl-dipeptida
7	11	35.5	7	2	A28709	phosphonoacetaldeh
8	11	35.5	7	2	E48394	glycoprotein compo
9	11	35.5	7	2	I48086	DNA topoisomerase
10	11	35.5	7	2	B48394	major fat-globule
11	10	32.3	4	2	I40505	hypothetical prote
12	10	32.3	5	2	T14910	hypothetical prote
13	10	32.3	6	2	I48126	alpha-tubulin - Ch
14	10	32.3	6	2	B33932	Ig mu chain D regi
15	10	32.3	7	2	ECMUCR	catch-relaxing pep
16	10	32.3	7	2	A30812	sex pheromone cCF1
17	10	32.3	7	2	A28340	myomodulin - Calif
18	10	32.3	7	2	PN0649	pullulanase (EC 3.
19	9	29.0	5	2	C41225	copper resistance
20	9	29.0	5	2	PT0525	T-cell receptor be
21	9	29.0	5	2	PT0577	T-cell receptor be
22	9	29.0	5	2	PT0565	T-cell receptor be
23	9	29.0	5	2	PT0700	T-cell receptor be
24	9	29.0	5	2	S69237	surface protein te
25	9	29.0	6	2	B34835	dnaA protein - Pse
26	9	29.0	6	2	A43766	28K ubiquitin-immu
27	9	29.0	6	2	I65546	MHC H2-L antigen -
28	9	29.0	6	2	PT0518	T-cell receptor be
29	9	29.0	6	2	PT0662	T-cell receptor be
30	9	29.0	7	2	A15398	choline oxidase (E
31	9	29.0	7	2	S25266	pilE protein - Esc
32	9	29.0	7	2	PN0150	omega-gliadine 1'
33	9	29.0	7	2	S78024	ribosomal protein
34	9	29.0	7	2	E30608	Ig kappa chain V-I
35	9	29.0	7	2	PT0671	T-cell receptor be
36	9	29.0	7	2	PX0008	glucuronosyltransf
37	9	29.0	7	2	S66442	glutathione S-tran
38	9	29.0	7	2	S09066	globulin IV alpha
39	8	25.8	4	2	I61883	protamine P1 - ora
40	8	25.8	4	2	I37013	protamine P1 - Cer
41	8	25.8	4	2	I84439	protamine P1 - sav
42	8	25.8	5	2	I39964	ribosomal protein
43	8	25.8	5	2	I39966	ribosomal protein
44	8	25.8	5	2	I39965	ribosomal protein
45	8	25.8	5	2	G44817	27.5 kda structura

ALIGNMENTS

RESULT 1
S19630
ribosomal protein L30 - Streptomyces griseus (fragment)
C;Species: Streptomyces griseus
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Jun-1997
C;Accession: S19630
R;Ochi, K.
Int. J. Syst. Bacteriol. 42, 144-150, 1992
A;Title: Electrophoretic heterogeneity of ribosomal protein AT-L30 among actinomycete
A;Reference number: S19630; MUID:92144363; PMID:1736962
A;Accession: S19630
A;Molecule type: protein
A;Residues: 1-7 <OCH>
A;Experimental source: strain IFO 13189
C;Superfamily: Escherichia coli ribosomal protein L30
C;Keywords: protein biosynthesis; ribosome

Query Match 41.9%; Score 13; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 LKLSR 7
              ||:::
Db          3 LKITQ 7
```

RESULT 2
T13892
cytochrome-c oxidase (EC 1.9.3.1) chain I [imported] - river lamprey mitochondrion (f
C;Species: mitochondrion Lampetra fluviatilis (river lamprey)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: T13892
R;Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A;Title: The main features of the craniate mitochondrial DNA between the ND1 and the
A;Reference number: Z17775; MUID:97398704; PMID:9254918
A;Accession: T13892
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3 <DEL>
A;Cross-references: EMBL:Y09528; NID:g2340016; PIDN:CAA70721.1; PID:g4379123
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

Query Match 35.5%; Score 11; DB 3; Length 3;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LTL 3
              :||
```

```
Db          1 MTL 3


RESULT 3
T46627
hypothetical protein c4 - loblolly pine
C;Species: Pinus taeda (loblolly pine)
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
C;Accession: T46627
R;Chang, S.; Puryea, J.; Funkhouser, E.A.; Newton, R.J.; Cairney, J.
submitted to the EMBL Data Library, July 1995
A;Description: Cloning of a chitinase homolog which lacks chitin binding sites and is dc
A;Reference number: Z23105
A;Accession: T46627
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-4 <CHA>
A;Cross-references: EMBL:U31309; NID:g974285; PID:g974292
A;Experimental source: strain s6PT2xs6PT3; 8 month seedlings

  Query Match             35.5%;  Score 11;  DB 2;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LKL 5
              :||
Db          1 MKL 3


RESULT 4
E42364
flagellar protein fliR - Salmonella typhimurium (fragment)
C;Species: Salmonella typhimurium
C;Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 30-Sep-1993
C;Accession: E42364
R;Vogler, A.P.; Homma, M.; Irikura, V.M.; Macnab, R.M.
J. Bacteriol. 173, 3564-3572, 1991
A;Title: Salmonella typhimurium mutants defective in flagellar filament regrowth and seq
A;Reference number: A42364; MUID:91258342; PMID:1646201
A;Accession: E42364
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-5 <VOG>
A;Cross-references: GB:M62408

  Query Match             35.5%;  Score 11;  DB 2;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LTL 3
              :||
Db          2 ITL 4


RESULT 5
I49424
cytotoxic T-lymphocyte proteinase 3 (EC 3.4.21.-) - western wild mouse (fragment)
C;Species: Mus spretus (western wild mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I49424
R;Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H
Mamm. Genome 5, 349-355, 1994
A;Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
A;Reference number: I48934; MUID:94319082; PMID:8043949
A;Accession: I49424
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-6 <RES>
A;Cross-references: EMBL:U05745; NID:g497084; PIDN:AAB60481.1; PID:g642831
C;Keywords: hydrolase; serine proteinase

  Query Match             35.5%;  Score 11;  DB 2;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LKL 5
              :||
Db          3 MKL 5


RESULT 6
JN0859
peptidyl-dipeptidase A inhibitory peptide C105 - striped bonito
C;Species: Sarda orientalis (striped bonito)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999
C;Accession: JN0859
R;Matsumura, N.; Fujii, M.; Takeda, Y.; Shimizu, T.
Biosci. Biotechnol. Biochem. 57, 1743-1744, 1993
A;Title: Isolation and characterization of angiotensin I-converting enzyme inhibitory
A;Reference number: JN0859; MUID:94080036; PMID:7764272
A;Accession: JN0859
A;Molecule type: protein
A;Residues: 1-7 <MAT>
A;Experimental source: intestine
C;Comment: The carboxyl-terminus is essential for the protein's expression of angioter
C;Superfamily: bradykinin-potentiating peptide
C;Keywords: angiotensin-converting enzyme inhibitor

  Query Match             35.5%;  Score 11;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 KLSR 7
              || :
Db          4 KLEK 7


RESULT 7
A28709
phosphonoacetaldehyde hydrolase - Bacillus cereus (fragment)
C;Species: Bacillus cereus
C;Date: 22-Aug-1988 #sequence_revision 22-Aug-1988 #text_change 30-Sep-1993
C;Accession: A28709
R;Olsen, D.B.; Hepburn, T.W.; Moos, M.; Mariano, P.S.; Dunaway-Mariano, D.
Biochemistry 27, 2229-2234, 1988
A;Title: Investigation of the Bacillus cereus phosphonoacetaldehyde hydrolase. Evidenc
idue.
A;Reference number: A28709; MUID:88241058; PMID:3132206
A;Accession: A28709
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <OLS>

  Query Match             35.5%;  Score 11;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LKL 5
              ||:
Db          1 LKI 3


RESULT 8
E48394
glycoprotein component 16/major fat-globule membrane protein/MFG-E8 homolog - bovine (
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 07-Feb-1997
C;Accession: E48394
R;Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A;Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-p
II-like sequences.
A;Reference number: A48394; MUID:93250576; PMID:8485470
A;Accession: E48394
A;Status: preliminary
A;Molecule type: protein
```

A;Residues: 1-7 <MAT>
A;Experimental source: milk
A;Note: sequence extracted from NCBI backbone (NCBIP:131450)
C;Keywords: glycoprotein

Query Match 35.5%; Score 11; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KLSR 7
:|:|
Db 4 ELAR 7

RESULT 9
I48086
DNA topoisomerase II alpha - Chinese hamster (fragment)
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I48086
R;Ng, S.W.; Eder, J.P.; Schnipper, L.E.; Chan, V.T.W.
J. Biol. Chem. 270, 25850-25858, 1995
A;Title: Molecular cloning and characterization of the promoter for the Chinese hamster
A;Reference number: I48086; MUID:96029684; PMID:7592770
A;Accession: I48086
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-7 <RES>
A;Cross-references: EMBL:U34196; NID:g1041231; PIDN:AAC52315.1; PID:g1041232

Query Match 35.5%; Score 11; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LKLS 6
::||
Db 1 MELS 4

RESULT 10
B48394
major fat-globule membrane protein GP 55 - guinea pig (fragment)
C;Species: Cavia porcellus (guinea pig)
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995
C;Accession: B48394
R;Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A;Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.
A;Reference number: A48394; MUID:93250576; PMID:8485470
A;Accession: B48394
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MAT>
A;Experimental source: milk
A;Note: sequence extracted from NCBI backbone (NCBIP:131444)

Query Match 35.5%; Score 11; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KLSR 7
:|:|
Db 4 ELAR 7

RESULT 11
I40505
hypothetical protein 3 (4 aa) - Bacillus stearothermophilus
C;Species: Bacillus stearothermophilus
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Oct-1999
C;Accession: I40505
R;Waye, M.M.; Winter, G.
Eur. J. Biochem. 158, 505-510, 1986
A;Title: A transcription terminator in the 5' non-coding region of the tyrosyl tRNA s
A;Reference number: I40503; MUID:86274732; PMID:3525162
A;Accession: I40505
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4 <RES>
A;Cross-references: EMBL:X04193; NID:g40233; PIDN:CAA27783.1; PID:g580944

Query Match 32.3%; Score 10; DB 2; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LSR 7
||:
Db 2 LSK 4

RESULT 12
T14910
hypothetical protein - parsley
C;Species: Petroselinum crispum (parsley)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14910
R;Kircher, S.; Ledger, S.; Hayashi, H.; Weisshaar, B.; Schafer, E.; Frohnmeyer, H.
Mol. Gen. Genet. 257, 595-605, 1998
A;Title: CPRF4a, a novel plant bZIP protein of the CPRF family: comparative analysis
A;Reference number: Z18261; MUID:98265918; PMID:9604882
A;Accession: T14910
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-5 <KIR>
A;Cross-references: EMBL:Y10810; NID:g3336904; PIDN:CAA71769.1; PID:g3336905
A;Experimental source: ssp. Hamburger Schnitt

Query Match 32.3%; Score 10; DB 2; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LSR 7
:||
Db 2 VSR 4

RESULT 13
I48126
alpha-tubulin - Chinese hamster (fragment)
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 05-Nov-1999
C;Accession: I48126
R;Elliott, E.M.; Sarangi, F.; Henderson, G.; Ling, V.
Can. J. Biochem. Cell Biol. 63, 511-518, 1985
A;Title: Cloning of 11 alpha-tubulin gene sequences from the genome of Chinese hamste
A;Reference number: I48126; MUID:86001952; PMID:2931165
A;Accession: I48126
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-6 <RES>
A;Cross-references: GB:M25895; NID:g341417; PIDN:AAA74493.1; PID:g516601
C;Genetics:
A;Introns: 3/3

Query Match 32.3%; Score 10; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KLS 6
||:
Db 2 KLA 4

RESULT 14
B33932

```
Ig mu chain D region (D23) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 16-Aug-1996
C;Accession: B33932
R;Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.
Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989
A;Title: Two murine natural polyreactive autoantibodies are encoded by nonmutated germ-l
A;Reference number: A33932; MUID:89282823; PMID:2499887
A;Accession: B33932
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-6 <BAC>
A;Cross-references: GB:M27107
C;Keywords: immunoglobulin

  Query Match             32.3%;  Score 10;  DB 2;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LKL 5
              |:|
Db          3 LRL 5


RESULT 15
ECMUCR
catch-relaxing peptide - blue mussel
N;Alternate names: CARP
C;Species: Mytilus edulis (blue mussel)
C;Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 20-Jun-2000
C;Accession: A29342
R;Hirata, T.; Kubota, I.; Takabatake, I.; Kawahara, A.; Shimamoto, N.; Muneoka, Y.
Brain Res. 422, 374-376, 1987
A;Title: Catch-relaxing peptide isolated from Mytilus pedal ganglia.
A;Reference number: A29342; MUID:88052022; PMID:3676797
A;Accession: A29342
A;Molecule type: protein
A;Residues: 1-7 <HIR>
C;Comment: This peptide exhibits both potentiating (contraction) and inhibitory (relaxat
C;Superfamily: unassigned animal peptides
C;Keywords: amidated carboxyl end; hormone; retractor muscle
F;7/Modified site: amidated carboxyl end (Leu) #status experimental

  Query Match             32.3%;  Score 10;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LKL 5
              |:|
Db          5 LRL 7


Search completed: June  2, 2004, 19:00:24
Job time : 14 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:47:58 ; Search time 10 Seconds
(without alignments)
36.449 Million cell updates/sec

Title: US-09-712-819D-13
Perfect score: 31
Sequence: 1 LTLKLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 88

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	10	32.3	7	1	CARP_MYTED	P10420	mytilus edu
2	10	32.3	7	1	CCF1_ENTFA	P20104	enterococcu
3	9	29.0	7	1	CHOX_ALCSP	P16101	alcaligenes
4	9	29.0	7	1	UC24_MAIZE	P80630	zea mays (m
5	8	25.8	5	1	EI04_LITRU	P82100	litoria rub
6	8	25.8	7	1	FAR5_HIRME	P42564	hirudo medi
7	8	25.8	7	1	GFRP_MOUSE	P99025	mus musculu
8	7	22.6	3	1	LUXE_VIBFI	P24272	vibrio fisc
9	7	22.6	5	1	BIOA_CITFR	P13071	citrobacter
10	7	22.6	5	1	UF01_MOUSE	P38639	mus musculu
11	7	22.6	6	1	ACPH_RABIT	P25154	oryctolagus
12	7	22.6	6	1	UN06_CLOPA	P81351	clostridium
13	7	22.6	6	1	VP19_HSV1K	P23210	herpes simp
14	7	22.6	7	1	ALL7_CYDPO	P82158	cydia pomon
15	7	22.6	7	1	WWA1_ACHFU	P35919	achatina fu
16	7	22.6	7	1	WWA2_ACHFU	P35920	achatina fu
17	7	22.6	7	1	WWA3_ACHFU	P35921	achatina fu
18	6	19.4	4	1	FAR3_HIRME	P42562	hirudo medi
19	6	19.4	4	1	FLRF_HIRME	P42561	hirudo medi
20	6	19.4	4	1	FLRN_ANTEL	P58707	anthopleura
21	6	19.4	4	1	RM01_YEAST	P36515	saccharomyc
22	6	19.4	6	1	OVM_LEPDE	P42985	leptinotars
23	6	19.4	7	1	FAR1_HELTI	P41871	helisoma tr
24	6	19.4	7	1	FAR1_MACRS	P83274	macrobrachi
25	6	19.4	7	1	FAR1_PROCL	P38499	procambarus
26	6	19.4	7	1	FAR2_ASCSU	P31890	ascaris suu
27	6	19.4	7	1	FAR2_PROCL	P38498	procambarus
28	6	19.4	7	1	UH11_RAT	P56576	rattus norv
29	5	16.1	3	1	GRWM_HUMAN	P01157	homo sapien
30	5	16.1	4	1	DCMS_PSECH	P19918	pseudomonas
31	5	16.1	4	1	FAR4_HIRME	P42563	hirudo medi
32	5	16.1	4	1	FFKA_ANTEL	P58705	anthopleura
33	5	16.1	4	1	FMRF_MACNI	P01162	macrocallis
34	5	16.1	4	1	FYRI_ANTEL	P58706	anthopleura
35	5	16.1	4	1	TUFT_HUMAN	P01858	homo sapien
36	5	16.1	5	1	BPP7_BOTIN	P30425	bothrops in
37	5	16.1	5	1	FARP_ARTTR	P41853	artioposthi
38	5	16.1	5	1	PRCT_PERAM	P01373	periplaneta
39	5	16.1	5	1	PSK_DAUCA	P58261	daucus caro
40	5	16.1	5	1	RE31_LITRU	P82072	litoria rub
41	5	16.1	5	1	RE32_LITRU	P82073	litoria rub
42	5	16.1	5	1	TRM3_ECOLI	P13973	escherichia
43	5	16.1	6	1	ASP2_LACSN	P82655	lactobacill
44	5	16.1	6	1	FARP_MONEX	P41966	moniezia ex
45	5	16.1	6	1	TMOF_SARBU	P41495	sarcophaga

ALIGNMENTS

RESULT 1
CARP_MYTED
ID CARP_MYTED STANDARD; PRT; 7 AA.
AC P10420;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE Catch-relaxing peptide (CARP).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloida;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RX MEDLINE=88052022; PubMed=3676797;
RA Hirata T., Kubota I., Takabatake I., Kawahara A., Shimamoto N.,
RA Muneoka Y.;
RT "Catch-relaxing peptide isolated from Mytilus pedal ganglia.";
RL Brain Res. 422:374-376(1987).
CC -!- FUNCTION: This peptide exhibits both potentiating (contraction)
CC and inhibitory (relaxation) effects on the anterior byssus
CC retractor muscle.
DR PIR; A29342; ECMUCR.
KW Hormone; Amidation.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 831 MW; 6734072687669DB0 CRC64;

Query Match 32.3%; Score 10; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 LKL 5
              |:|
Db          5 LRL 7
```

RESULT 2
CCF1_ENTFA
ID CCF1_ENTFA STANDARD; PRT; 7 AA.
AC P20104;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone cCF10.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RX MEDLINE=89008313; PubMed=3139658;
RA Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,
RA Adsit J.C., Dunny G.M., Suzuki A.;
RT "Structure of cCF10, a peptide sex pheromone which induces
RT conjugative transfer of the Streptococcus faecalis tetracycline
RT resistance plasmid, pCF10.";
RL J. Biol. Chem. 263:14574-14578(1988).

```
CC   -!- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       HEMOLYSIN PLASMID PCF10.
DR   PIR; A30812; A30812.
KW   Pheromone.
SQ   SEQUENCE   7 AA;  790 MW;  72C9D2C731B2C740 CRC64;

  Query Match             32.3%;  Score 10;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LTL 3
              :||
Db          2 VTL 4


RESULT 3
CHOX_ALCSP
ID   CHOX_ALCSP     STANDARD;      PRT;     7 AA.
AC   P16101;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-APR-1990 (Rel. 14, Last annotation update)
DE   Choline oxidase (EC 1.1.3.17) (Fragment).
OS   Alcaligenes sp.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Alcaligenes.
OX   NCBI_TaxID=512;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=81006769; PubMed=6997283;
RA   Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;
RT   "Identification and properties of the prosthetic group of choline
RT   oxidase from Alcaligenes sp.";
RL   J. Biochem. 88:197-203(1980).
CC   -!- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).
DR   PIR; A15398; A15398.
KW   Oxidoreductase.
FT   NON_TER       7       7
SQ   SEQUENCE   7 AA;  839 MW;  7415B1E457644AC0 CRC64;

  Query Match             29.0%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 SR 7
              ||
Db          6 SR 7


RESULT 4
UC24_MAIZE
ID   UC24_MAIZE     STANDARD;      PRT;     7 AA.
AC   P80630;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Unknown protein from 2D-page of etiolated coleoptile (Spot 447)
DE   (Fragment).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Coleoptile;
RA   Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA   Pernollet J.-C., Zivy M., de Vienne D.;
RT   "The maize two dimensional gel protein database: towards an integrated
RT   genome analysis program.";
RL   Theor. Appl. Genet. 93:997-1005(1996).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 6.0, its MW is: 30.0 kDa.
DR   Maize-2DPAGE; P80630; COLEOPTILE.
DR   MaizeDB; 123956; -.
FT   NON_TER       1       1
FT   NON_TER       7       7
SQ   SEQUENCE   7 AA;  665 MW;  6DC1B5B33DC1B5D0 CRC64;

  Query Match             29.0%;  Score 9;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLK 4
              | |
Db          2 TAK 4


RESULT 5
EI04_LITRU
ID   EI04_LITRU     STANDARD;      PRT;     5 AA.
AC   P82100;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Electrin 4.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Skin secretion;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litori electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:639-645(1999).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Skin.
KW   Amphibian defense peptide; Amidation.
FT   MOD_RES       5       5       AMIDATION.
SQ   SEQUENCE   5 AA;  616 MW;  61F2D1A059A00000 CRC64;

  Query Match             25.8%;  Score 8;  DB 1;  Length 5;
  Best Local Similarity   33.3%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LTL 3
              :|:
Db          2 ITV 4


RESULT 6
FAR5_HIRME
ID   FAR5_HIRME     STANDARD;      PRT;     7 AA.
AC   P42564;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFamide-like neuropeptide GGKYMRF-amide.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
KW   Neuropeptide; Amidation.
FT   MOD_RES       7       7       AMIDATION.
```

```
SQ   SEQUENCE   7 AA;  858 MW;  69D4068B53387810 CRC64;

  Query Match             25.8%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 KLSR 7
              | |
Db          3 KYMR 6


RESULT 7
GFRP_MOUSE
ID   GFRP_MOUSE     STANDARD;      PRT;     7 AA.
AC   P99025;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
GN   GCHFR OR GFRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Liver;
RA   Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
RA   Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA   Cowthorne M.;
RL   Submitted (AUG-1998) to Swiss-Prot.
CC   -!- FUNCTION: Mediates tetrahydrobiopterin inhibition of GTP
CC       cyclohydrolase I. This inhibition is reversed by L-phenylalanine
CC       (By similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
DR   SWISS-2DPAGE; P99025; MOUSE.
FT   INIT_MET       0       0
FT   NON_TER        7       7
SQ   SEQUENCE   7 AA;  806 MW;  71B5B057273B4700 CRC64;

  Query Match             25.8%;  Score 8;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LKLS 6
              | :|
Db          3 LLIS 6


RESULT 8
LUXE_VIBFI
ID   LUXE_VIBFI     STANDARD;      PRT;     3 AA.
AC   P24272;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Long-chain-fatty-acid--luciferin-component ligase (EC 6.2.1.19) (Acyl-
DE   protein synthetase) (Fragment).
GN   LUXE.
OS   Vibrio fischeri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=668;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91072226; PubMed=2254256;
RA   Swartzman E., Kapoor S., Graham A.F., Meighen E.A.;
RT   "A new Vibrio fischeri lux gene precedes a bidirectional termination
RT   site for the lux operon.";
RL   J. Bacteriol. 172:6797-6802(1990).
CC   -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.
CC       IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE
CC       FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS
CC       SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.
CC   -!- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +
CC       an acyl-protein thiolester.
CC   -!- PATHWAY: Bioluminescent fatty acid reduction system; second step.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M62812; -; NOT_ANNOTATED_CDS.
KW   Luminescence; Ligase.
FT   NON_TER        1       1
SQ   SEQUENCE   3 AA;  374 MW;  6AA3303000000000 CRC64;

  Query Match             22.6%;  Score 7;  DB 1;  Length 3;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
              :|
Db          1 IK 2


RESULT 9
BIOA_CITFR
ID   BIOA_CITFR     STANDARD;      PRT;     5 AA.
AC   P13071;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
DE   (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA
DE   aminotransferase) (Fragment).
GN   BIOA.
OS   Citrobacter freundii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Citrobacter.
OX   NCBI_TaxID=546;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89006280; PubMed=2971595;
RA   Shiuan D., Campbell A.;
RT   "Transcriptional regulation and gene arrangement of Escherichia coli,
RT   Citrobacter freundii and Salmonella typhimurium biotin operons.";
RL   Gene 67:203-211(1988).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-
CC       oxononanoate = S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-
CC       diaminononanoate.
CC   -!- COFACTOR: Pyridoxal phosphate.
CC   -!- PATHWAY: Biotin biosynthesis.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SIMILARITY: Belongs to class-III of pyridoxal-phosphate-dependent
CC       aminotransferases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21922; -; NOT_ANNOTATED_CDS.
DR   PIR; I40697; I40697.
DR   InterPro; IPR005814; Aminotrans_3.
DR   PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.
KW   Biotin biosynthesis; Transferase; Aminotransferase;
KW   Pyridoxal phosphate.
FT   NON_TER        5       5
```

```
SQ   SEQUENCE   5 AA;  582 MW;  6AAAB1B1A6F00000 CRC64;

  Query Match             22.6%;  Score 7;  DB 1;  Length 5;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LT 2
              :|
Db          1 MT 2


RESULT 10
UF01_MOUSE
ID   UF01_MOUSE     STANDARD;      PRT;     5 AA.
AC   P38639;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Unknown protein from 2D-page of fibroblasts (P19) (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Fibroblast;
RX   MEDLINE=95009907; PubMed=7523108;
RA   Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT   "Separation and sequencing of familiar and novel murine proteins
RT   using preparative two-dimensional gel electrophoresis.";
RL   Electrophoresis 15:735-745(1994).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 6.6, its MW is: 19 kDa.
FT   NON_TER       5       5
SQ   SEQUENCE   5 AA;  717 MW;  7364087043100000 CRC64;

  Query Match             22.6%;  Score 7;  DB 1;  Length 5;
  Best Local Similarity   33.3%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 LSR 7
              : |
Db          2 IGR 4


RESULT 11
ACPH_RABIT
ID   ACPH_RABIT     STANDARD;      PRT;     6 AA.
AC   P25154;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Acylamino-acid-releasing enzyme (EC 3.4.19.1) (AARE) (Acyl-peptide
DE   hydrolase) (APH) (Acylaminoacyl-peptidase) (Fragment).
GN   APEH.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Muscle;
RX   MEDLINE=92222120; PubMed=1807161;
RA   Krishna R.G., Chin C.C.Q., Wold F.;
RT   "N-terminal sequence analysis of N alpha-acetylated proteins after
RT   unblocking with N-acylaminoacyl-peptide hydrolase.";
RL   Anal. Biochem. 199:45-50(1991).
CC   -!- FUNCTION: This enzyme catalyzes the hydrolysis of the N-terminal
CC       peptide bond of an N-acetylated peptide to generate an N-
CC       acetylated amino acid and a peptide with a free N-terminus. It
CC       preferentially cleaves off Ac-Ala, Ac-Met and Ac-Ser.
CC   -!- CATALYTIC ACTIVITY: Acylaminoacyl-peptide + H(2)O = acylamino acid
CC       + peptide.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to peptidase family S9C.
DR   PIR; A49792; A49792.
DR   MEROPS; S09.004; -.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; PARTIAL.
KW   Hydrolase; Acetylation.
FT   MOD_RES       1       1       ACETYLATION.
FT   NON_TER       6       6
SQ   SEQUENCE   6 AA;  775 MW;  6732D6C40B16F000 CRC64;

  Query Match             22.6%;  Score 7;  DB 1;  Length 6;
  Best Local Similarity   33.3%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 LSR 7
              : |
Db          1 MER 3


RESULT 12
UN06_CLOPA
ID   UN06_CLOPA     STANDARD;      PRT;     6 AA.
AC   P81351;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Unknown protein CP 6 from 2D-page (Fragment).
OS   Clostridium pasteurianum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1501;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=W5;
RX   MEDLINE=98291870; PubMed=9629918;
RA   Flengsrud R., Skjeldal L.;
RT   "Two-dimensional gel electrophoresis separation and N-terminal
RT   sequence analysis of proteins from Clostridium pasteurianum W5.";
RL   Electrophoresis 19:802-806(1998).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 5.0, ITS MW IS: 75.9 kDa.
FT   NON_TER       6       6
SQ   SEQUENCE   6 AA;  657 MW;  605B1DC1A45A8000 CRC64;

  Query Match             22.6%;  Score 7;  DB 1;  Length 6;
  Best Local Similarity   25.0%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKL 5
              | ::
Db          3 TAEI 6


RESULT 13
VP19_HSV1K
ID   VP19_HSV1K     STANDARD;      PRT;     6 AA.
AC   P23210;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Capsid assembly and DNA maturation protein (Virion protein UL38)
DE   (Capsid protein VP19C) (Fragment).
GN   UL38.
OS   Herpes simplex virus (type 1 / strain KOS).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=10306;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91101287; PubMed=1846198;
RA   Flanagan W.M., Papavassiliou A.G., Rice M., Hecht L.B.,
```

```
RA   Silverstein S., Wagner E.K.;
RT   "Analysis of the herpes simplex virus type 1 promoter controlling the
RT   expression of UL38, a true late gene involved in capsid assembly.";
RL   J. Virol. 65:769-786(1991).
CC   -!- FUNCTION: COMPONENT OF THE BASAL LAYER IN WHICH THE CAPSIDS ARE
CC       EMBEDDED. BINDS DNA.
CC   -!- SIMILARITY: Belongs to the herpesviruses VP19C family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57646; AAA45830.1; -.
KW   Capsid assembly; Coat protein; DNA-binding.
FT   NON_TER       6       6
SQ   SEQUENCE   6 AA;  703 MW;  67376451A336F000 CRC64;

  Query Match             22.6%;  Score 7;  DB 1;  Length 6;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
              :|
Db          1 MK 2


RESULT 14
ALL7_CYDPO
ID   ALL7_CYDPO     STANDARD;      PRT;     7 AA.
AC   P82158;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Cydiastatin 7.
OS   Cydia pomonella (Codling moth).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Tortricoidea; Tortricidae; Olethreutinae; Cydia.
OX   NCBI_TaxID=82600;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Larva;
RX   MEDLINE=98054539; PubMed=9392829;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA   Davey M., East P.D., Thorpe A.;
RT   "Lepidopteran peptides of the allatostatin superfamily.";
RL   Peptides 18:1301-1309(1997).
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Neuropeptide; Amidation.
FT   MOD_RES       7       7       AMIDATION.
SQ   SEQUENCE   7 AA;  873 MW;  672879CABB569350 CRC64;

  Query Match             22.6%;  Score 7;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 KL 5
              |:
Db          1 KM 2


RESULT 15
WWA1_ACHFU
ID   WWA1_ACHFU     STANDARD;      PRT;     7 AA.
AC   P35919;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   WWamide-1.
OS   Achatina fulica (Giant African snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC   Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX   NCBI_TaxID=6530;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Ganglion;
RX   MEDLINE=93265912; PubMed=8495720;
RA   Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT   "WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT   ganglia of the African giant snail, Achatina fulica.";
RL   FEBS Lett. 323:104-108(1993).
CC   -!- FUNCTION: Exhibits modulatory effects on the peripheral nervous
CC       system. Inhibits activity on a central neuron.
DR   PIR; S33245; S33245.
KW   Neuropeptide; Amidation.
FT   MOD_RES       7       7       AMIDATION.
SQ   SEQUENCE   7 AA;  993 MW;  7362D5B69B041310 CRC64;

  Query Match             22.6%;  Score 7;  DB 1;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 1.4e+05;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 KLS 6
              ::|
Db          3 EMS 5


Search completed: June  2, 2004, 18:58:23
Job time : 11 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:48:29 ; Search time 34.5 Seconds
(without alignments)
64.018 Million cell updates/sec

Title: US-09-712-819D-13
Perfect score: 31
Sequence: 1 LTLKLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 74

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	13	41.9	7	8	Q95945	Q95945 saccharomyc
2	11	35.5	6	10	P82181	P82181 spinacia ol
3	11	35.5	6	10	P82182	P82182 spinacia ol
4	11	35.5	7	2	Q8GL00	Q8gl00 borrelia bu
5	11	35.5	7	15	Q8JE81	Q8je81 human immun
6	10	32.3	6	10	P82541	P82541 spinacia ol
7	10	32.3	7	2	P83530	P83530 lactobacill
8	9	29.0	4	11	Q08433	Q08433 rattus sp.
9	9	29.0	5	13	P83308	P83308 gallus gall
10	9	29.0	7	2	O07354	O07354 synechococc
11	9	29.0	7	2	Q8GL12	Q8gl12 borrelia bu
12	9	29.0	7	2	Q8GL04	Q8gl04 borrelia bu
13	9	29.0	7	4	Q15897	Q15897 homo sapien
14	9	29.0	7	5	Q7Z1C0	Q7z1c0 caenorhabdi
15	9	29.0	7	10	P93233	P93233 lycopersico
16	9	29.0	7	11	Q8K3H6	Q8k3h6 rattus norv
17	9	29.0	7	12	Q66205	Q66205 transmissib
18	8	25.8	7	2	P70804	P70804 azotobacter
19	8	25.8	7	2	O34028	O34028 sphingomona
20	8	25.8	7	10	Q9C5B3	Q9c5b3 arabidopsis
21	8	25.8	7	12	Q9YVE3	Q9yve3 human adeno
22	8	25.8	7	12	Q66113	Q66113 cherry leaf
23	8	25.8	7	12	Q9YIQ9	Q9yiq9 human adeno
24	8	25.8	7	12	Q9YIR0	Q9yir0 human adeno
25	8	25.8	7	13	O42564	O42564 fugu rubrip
26	7	22.6	5	2	P83073	P83073 bacillus ce
27	7	22.6	7	2	Q47029	Q47029 enterobacte
28	7	22.6	7	2	P72081	P72081 nocardia la
29	7	22.6	7	2	Q54248	Q54248 streptomyce
30	7	22.6	7	6	Q28742	Q28742 oryctolagus
31	7	22.6	7	8	P92214	P92214 amblyopyrum
32	7	22.6	7	8	P92393	P92393 hordeum vul
33	7	22.6	7	8	P92403	P92403 lophopyrum
34	7	22.6	7	8	P92427	P92427 peridictyon
35	7	22.6	7	8	P92430	P92430 aegilops ta
36	7	22.6	7	8	P92221	P92221 bromus iner
37	7	22.6	7	8	P92425	P92425 pseudoroegn
38	7	22.6	7	8	P92381	P92381 hordeum bra
39	7	22.6	7	8	P92387	P92387 henrardia p
40	7	22.6	7	8	P92210	P92210 agropyron c
41	7	22.6	7	8	P92440	P92440 thinopyrum
42	7	22.6	7	8	P92218	P92218 australopyr
43	7	22.6	7	8	P92390	P92390 heteranthel
44	7	22.6	7	8	P92372	P92372 haynaldia v
45	7	22.6	7	8	P92442	P92442 taeniatheru

ALIGNMENTS

RESULT 1
Q95945
ID Q95945 PRELIMINARY; PRT; 7 AA.
AC Q95945;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Inside intron 5 (Fragment).
OS Saccharomyces cerevisiae (Baker's yeast).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D273-10B;
RX MEDLINE=81069885; PubMed=6254986;
RA Bonitz S.G., Coruzzi G., Thalenfeld B., Tzagoloff A., Macino G.;
RT "Assembly of the mitochondrial membrane system: Structure and
RT nucleotide sequence of the gene coding for subunit 1 of yeast
RT cytochrome oxidase.";
RL J. Biol. Chem. 255:11927-11941(1980).
DR EMBL; V00694; CAA24066.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 859 MW; 75B7232362CDC460 CRC64;

Query Match 41.9%; Score 13; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          4 KLS 6
              |||
Db          5 KLS 7
```

RESULT 2
P82181

```
ID   P82181      PRELIMINARY;      PRT;     6 AA.
AC   P82181;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 50S ribosomal protein L10 beta (Fragment).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435798; PubMed=10874046;
RA   Yamaguchi K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the 50 S subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 275:28466-28482(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC   -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002363; Ribosomal_L10eub.
DR   PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  675 MW;  6321B415B05DB000 CRC64;

  Query Match             35.5%;  Score 11;  DB 10;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LSR 7
              :||
Db          2 ISR 4


RESULT 3
P82182
ID   P82182      PRELIMINARY;      PRT;     6 AA.
AC   P82182;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 50S ribosomal protein L10 gamma (Fragment).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435798; PubMed=10874046;
RA   Yamaguchi K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the 50 S subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 275:28466-28482(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC   -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002363; Ribosomal_L10eub.
DR   PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  675 MW;  6321B415B05DB000 CRC64;

  Query Match             35.5%;  Score 11;  DB 10;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LSR 7
              :||
Db          2 ISR 4


RESULT 4
Q8GL00
ID   Q8GL00      PRELIMINARY;      PRT;     7 AA.
AC   Q8GL00;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   PF-50 protein (Fragment).
GN   PF-50.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid group cp32-13.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CA15;
RA   Stevenson B., Miller J.C.;
RT   "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
RT   prophages: conservation amidst diversity.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY142106; AAN17857.1; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW   Plasmid.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  928 MW;  6337233050437350 CRC64;

  Query Match             35.5%;  Score 11;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LKL 5
              :||
Db          4 IKL 6


RESULT 5
Q8JE81
ID   Q8JE81      PRELIMINARY;      PRT;     7 AA.
AC   Q8JE81;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Truncated pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=4874;
RX   MEDLINE=22056123; PubMed=12060770;
RA   Beerenwinkel N., Schmidt B., Walter H., Kaiser R., Lengauer T.,
RA   Hoffmann D., Korn K., Selbig J.;
RT   "Diversity and complexity of HIV-1 drug resistance: A bioinformatics
RT   approach to predicting phenotype from genotype.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:8271-8276(2002).
DR   EMBL; AF347267; AAK32344.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  885 MW;  76C37731A046C700 CRC64;
```

```
  Query Match             35.5%;  Score 11;  DB 15;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LTL 3
              :||
Db          3 ITL 5


RESULT 6
P82541
ID   P82541      PRELIMINARY;      PRT;     6 AA.
AC   P82541;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 30S ribosomal protein S19 beta (Fragment).
OS   Spinacia oleracea (Spinach).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435797; PubMed=10874039;
RA   Yamaguchi K., von Knoblauch K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the small subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 37:28455-28465(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.
CC   -!- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.
CC   -!- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA
CC       FORM IS THE MINOR BASIC FORM.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.
CC   -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002222; Ribosomal_S19.
DR   PROSITE; PS00323; RIBOSOMAL_S19; PARTIAL.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  732 MW;  63333735A411C000 CRC64;

  Query Match             32.3%;  Score 10;  DB 10;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLK 4
              :||
Db          3 SLK 5


RESULT 7
P83530
ID   P83530      PRELIMINARY;      PRT;     7 AA.
AC   P83530;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Unknown protein from 2D-page (Fragment).
OS   Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1625;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=DSM 20451;
RX   PubMed=12112860;
RA   Drews O., Weiss W., Reil G., Parlar H., Wait R., Goerg A.;
RT   "High pressure effects step-wise altered protein expression in
RT   Lactobacillus sanfranciscensis.";
RL   Proteomics 2:765-774(2002).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS UNKNOWN
CC       PROTEIN IS: 15 KDA.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  790 MW;  6AAB02CAA731B2A0 CRC64;

  Query Match             32.3%;  Score 10;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LTL 3
              :||
Db          1 VTL 3


RESULT 8
Q08433
ID   Q08433      PRELIMINARY;      PRT;     4 AA.
AC   Q08433;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Bilirubin UDP-glucuronosyltransferase (Fragment).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Gunn;
RX   MEDLINE=91282758; PubMed=1840486;
RA   Sato H., Aono S., Kashiwamata S., Koiwai O.;
RT   "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
RT   hyperbilirubinemic Gunn rat.";
RL   Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
DR   EMBL; S38636; AAB19259.1; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
KW   Transferase.
FT   NON_TER       1      1
SQ   SEQUENCE   4 AA;  473 MW;  633732C420000000 CRC64;

  Query Match             29.0%;  Score 9;  DB 11;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
              ||
Db          3 LK 4


RESULT 9
P83308
ID   P83308      PRELIMINARY;      PRT;     5 AA.
AC   P83308;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   FMRFamide-like neuropeptide (LPLRF-amide).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Brain;
RX   PubMed=6137771;
RA   Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;
```

```
RT   "A novel active pentapeptide from chicken brain identified by
RT   antibodies to FMRFamide.";
RL   Nature 305:328-330(1983).
CC   -!- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   GO; GO:0007218; P:neuropeptide signaling pathway; TAS.
KW   Neuropeptide; Amidation.
FT   MOD_RES       5       5       AMIDATION.
SQ   SEQUENCE   5 AA;  645 MW;  69D4073767400000 CRC64;

  Query Match             29.0%;  Score 9;  DB 13;  Length 5;
  Best Local Similarity   50.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LTLK 4
              | |:
Db          1 LPLR 4


RESULT 10
O07354
ID   O07354      PRELIMINARY;      PRT;     7 AA.
AC   O07354;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NifK (Fragment).
GN   NIFK.
OS   Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothece PCC 8801).
OC   Bacteria; Cyanobacteria; Chroococcales; Cyanothece.
OX   NCBI_TaxID=41431;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RF-1;
RX   MEDLINE=99231861; PubMed=10217509;
RA   Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RT   "Organization and expression of nitrogen-fixation genes in the aerobic
RT   nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
RT   RF-1.";
RL   Microbiology 145:743-753(1999).
DR   EMBL; AF003700; AAC35193.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   7 AA;  849 MW;  7412C72AA9D5B030 CRC64;

  Query Match             29.0%;  Score 9;  DB 2;  Length 7;
  Best Local Similarity   28.6%;  Pred. No. 1e+06;
  Matches    2;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 LTLKLSR 7
              ::  | |
Db          1 ISFDLVR 7


RESULT 11
Q8GL12
ID   Q8GL12      PRELIMINARY;      PRT;     7 AA.
AC   Q8GL12;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   PF-50 protein (Fragment).
GN   PF-50.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid group cp32-9.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N40;
RA   Stevenson B., Miller J.C.;
RT   "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
RT   prophages: conservation amidst diversity.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY142100; AAN17911.1; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW   Plasmid.
FT   NON_TER       1       1
SQ   SEQUENCE   7 AA;  849 MW;  6337244330569ED0 CRC64;

  Query Match             29.0%;  Score 9;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
              ||
Db          6 LK 7


RESULT 12
Q8GL04
ID   Q8GL04      PRELIMINARY;      PRT;     7 AA.
AC   Q8GL04;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   PF-50 protein (Fragment).
GN   PF-50.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid group cp32-5.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=93-0107;
RA   Stevenson B., Miller J.C.;
RT   "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
RT   prophages: conservation amidst diversity.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY142103; AAN17848.1; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW   Plasmid.
FT   NON_TER       1       1
SQ   SEQUENCE   7 AA;  914 MW;  6337244330504310 CRC64;

  Query Match             29.0%;  Score 9;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
              ||
Db          6 LK 7


RESULT 13
Q15897
ID   Q15897      PRELIMINARY;      PRT;     7 AA.
AC   Q15897;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   (Clone XP6A11A) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA   Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA   Caskey C.T.H.;
RT   "Isolation of chromosome-specific genes by reciprocal probing of
RT   arrayed cDNAs and cosmid libraries.";
RL   Hum. Mol. Genet. 0:0-0(1995).
DR   EMBL; L32077; AAA73887.1; -.
```

```
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  814 MW;  672B1DD3372046B0 CRC64;

  Query Match             29.0%;  Score 9;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LK 4
              ||
Db          3 LK 4


RESULT 14
Q7Z1C0
ID   Q7Z1C0      PRELIMINARY;      PRT;     7 AA.
AC   Q7Z1C0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein W01B11.6.
GN   W01B11.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Wilson R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Bradshaw H., Graves T., Blair T.;
RT   "The sequence of C. elegans cosmid W01B11.";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Wilson R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF043704; AAO38592.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   7 AA;  874 MW;  72D1A9DB5041A6F0 CRC64;

  Query Match             29.0%;  Score 9;  DB 5;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 1e+06;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LTL 3
              :|:
Db          1 MTI 3


RESULT 15
P93233
ID   P93233      PRELIMINARY;      PRT;     7 AA.
AC   P93233;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14)
DE   (Fragment).
GN   LE-ACS1B.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97351561; PubMed=9207843;
RA   Oetiker J.H., Olson D.C., Shiu O.Y., Yang S.F.;
RT   "Differential induction of seven 1-aminocyclopropane-1-carboxylate
RT   synthase genes by elicitor in suspension cultures of tomato
RT   (Lycopersicon esculentum).";
RL   Plant Mol. Biol. 34:275-286(1997).
DR   EMBL; U75692; AAC49682.1; -.
DR   GO; GO:0016847; F:1-aminocyclopropane-1-carboxylate synthase . . .; IEA.
DR   GO; GO:0016829; F:lyase activity; IEA.
KW   Lyase.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  828 MW;  71B412C7377415D0 CRC64;

  Query Match             29.0%;  Score 9;  DB 10;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+06;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 SR 7
              ||
Db          1 SR 2


Search completed: June  2, 2004, 18:59:46
Job time : 35.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:47:08 ; Search time 49 Seconds
(without alignments)
40.364 Million cell updates/sec

Title: US-09-712-819D-13
Perfect score: 31
Sequence: 1 LTLKLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 92273

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	64.5	7	5	ABP66518	Abp66518 Human RSV
2	20	64.5	7	6	ABU69381	Abu69381 Respirato
3	20	64.5	7	7	ADE35876	Ade35876 SYNAGIS a
4	19	61.3	7	2	AAW69269	Aaw69269 Haemagglu
5	19	61.3	7	2	AAY42013	Aay42013 Rheumatoi
6	19	61.3	7	2	AAY41889	Aay41889 Rheumatoi
7	19	61.3	7	4	ABB55870	Abb55870 Vascular
8	19	61.3	7	4	ABB56283	Abb56283 Vascular
9	19	61.3	7	4	ABB55981	Abb55981 Vascular
10	19	61.3	7	4	AAU28602	Aau28602 DPI trypt
11	19	61.3	7	4	AAU24969	Aau24969 Schizophr
12	19	61.3	7	4	AAU26249	Aau26249 Depressio
13	19	61.3	7	4	AAU15313	Aau15313 Schizophr
14	19	61.3	7	4	ABB52355	Abb52355 Human API
15	19	61.3	7	5	ABP66489	Abp66489 Human RSV
16	19	61.3	7	5	ABG78901	Abg78901 Multiple
17	19	61.3	7	5	ABG78730	Abg78730 Multiple
18	19	61.3	7	6	ABP58010	Abp58010 Prostate
19	19	61.3	7	6	ABU69352	Abu69352 Respirato
20	19	61.3	7	6	ABP57255	Abp57255 Breast ca
21	19	61.3	7	6	ABR59042	Abr59042 Alzheimer
22	19	61.3	7	7	ADE35847	Ade35847 SYNAGIS a
23	18	58.1	6	5	AAE21359	Aae21359 S. cerevi
24	18	58.1	7	2	AAW82668	Aaw82668 Cauliflow
25	18	58.1	7	3	AAB02931	Aab02931 Nucleotid
26	18	58.1	7	4	AAE10492	Aae10492 Humanised
27	18	58.1	7	5	ABP66389	Abp66389 Human RSV
28	18	58.1	7	5	ABP66450	Abp66450 Human RSV
29	18	58.1	7	5	ABP66485	Abp66485 Human RSV
30	18	58.1	7	5	ABP49796	Abp49796 Zinc fing
31	18	58.1	7	5	ABP48455	Abp48455 Zinc fing
32	18	58.1	7	5	ABP51172	Abp51172 Zinc fing
33	18	58.1	7	5	ABP49577	Abp49577 Zinc fing
34	18	58.1	7	5	ABP49790	Abp49790 Zinc fing
35	18	58.1	7	5	ABB07139	Abb07139 Human ery
36	18	58.1	7	5	AAE28044	Aae28044 Human mod
37	18	58.1	7	6	ABU69348	Abu69348 Respirato
38	18	58.1	7	6	ABU69252	Abu69252 Respirato
39	18	58.1	7	6	ABU69313	Abu69313 Respirato
40	18	58.1	7	7	ADA65532	Ada65532 Zinc fing
41	18	58.1	7	7	ADA62425	Ada62425 Zinc fing
42	18	58.1	7	7	ADA64000	Ada64000 Zinc fing
43	18	58.1	7	7	ADA65571	Ada65571 Zinc fing
44	18	58.1	7	7	ADA65573	Ada65573 Zinc fing
45	18	58.1	7	7	ADE35808	Ade35808 SYNAGIS a

ALIGNMENTS

RESULT 1
ABP66518
ID ABP66518 standard; peptide; 7 AA.
XX
AC ABP66518;
XX
DT 04-DEC-2002 (first entry)
XX
DE Human RSV antibody VL CDR2 fragment.
XX
KW Human; variable heavy domain; variable light domain; CDR; VH; VL; RSV;
KW complementarity determining region; respiratory syncytial virus;
KW virucide; pulmonary; antiinflammatory; cardiant; anti-HIV; vaccine;
KW immunostimulant; gene therapy; cystic fibrosis; bone marrow transplant;
KW bronchopulmonary dysplasia; congenital heart disease;
KW congenital immunodeficiency; acquired immunodeficiency.
XX
OS Homo sapiens.
XX
PN WO200243660-A2.
XX
PD 06-JUN-2002.
XX
PF 28-NOV-2001; 2001WO-US044807.
XX
PR 28-NOV-2000; 2000US-00724396.
PR 28-NOV-2000; 2000US-00724531.
XX
PA (MEDI-) MEDIUMMUNE INC.
XX
PI Young JF, Koenig S, Johnson LS;
XX
DR WPI; 2002-706803/76.
XX
PT Antibody for treating respiratory syncytial virus (RSV) infection,
PT comprises a variable heavy/light domain or complementarity determining
PT regions 1 - 3 of variable light/heavy chains, that immunospecifically
PT binds to RSV antigen.
XX
PS Claim 8; Page 55; 298pp; English.
XX
CC The invention relates to a novel antibody comprising a variable heavy
CC (VH) domain, variable light (VL) domain, VH complementarity determining
CC region (CDR)-1, VH CDR2, VH CDR3, VL CDR1, VL CDR2 or VL CDR3, where the
CC antibody immunospecifically binds to a respiratory syncytial virus (RSV)
CC antigen, and where the antibody is not SYNAGIS (RTM). The antibody of the
CC invention has virucide, pulmonary, antiinflammatory, cardiant, anti-HIV,
CC and immunostimulant activity. The polynucleotides of the invention may

CC have a use in a vaccine, and in gene therapy. The antibody is useful for
CC treating or ameliorating a RSV infection in a human. The antibody is also
CC useful for preventing, treating or ameliorating one or more symptoms
CC associated with RSV infection in a mammal, e.g. cystic fibrosis,
CC bronchopulmonary dysplasia, congenital heart disease, congenital
CC immunodeficiency or acquired immunodeficiency, or after a bone marrow
CC transplant. The sequence represents a complementary determining region
CC peptide from a human RSV antibody of the invention
XX
SQ Sequence 7 AA;

Query Match 64.5%; Score 20; DB 5; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLKLS 6
|:|||
Db 2 TMKLS 6

RESULT 2
ABU69381
ID ABU69381 standard; peptide; 7 AA.
XX
AC ABU69381;
XX
DT 03-JUN-2003 (first entry)
XX
DE Respiratory syncytial virus (RSV) antibody fragment #126.
XX
KW Respiratory syncytial virus; RSV; vaccine; antibody;
KW variable heavy domain; VH; variable light domain; VL;
KW complementarity determining region; CDR; CDR1; CDR2; CDR3; RSV infection;
KW cystic fibrosis; bronchopulmonary dysplasia; congenital heart disease;
KW congenital immunodeficiency; acquired immune deficiency;
KW bone marrow transplant.
XX
OS Homo sapiens.
XX
PN US2002177126-A1.
XX
PD 28-NOV-2002.
XX
PF 28-NOV-2001; 2001US-00996288.
XX
PR 28-NOV-2000; 2000US-00724531.
XX
PA (MEDI-) MEDIMMUNE INC.
XX
PI Young JF, Koenig S, Johnson LS;
XX
DR WPI; 2003-340947/32.
XX
PT New antibody comprising a variable heavy (VH) or variable light (VL)
PT domain or complementarity determining region (CDR), such as CDR1, CDR2,
PT or CDR3, useful for preventing or treating a respiratory syncytial virus
PT (RSV) infection.
XX
PS Claim 8; Page 24; 165pp; English.
XX
CC The invention describes an antibody comprising a variable heavy (VH) or
CC variable light (VL) domain or complementarity determining region (CDR),
CC such as CDR1, CDR2, or CDR3, which immunospecifically binds to a
CC respiratory syncytial virus (RSV) antigen and is not SYNAGIS (RTM). The
CC antibody, pharmaceutical compositions and methods are useful for
CC preventing, treating or ameliorating a RSV infection in patients with
CC cystic fibrosis, bronchopulmonary dysplasia, congenital heart disease,
CC congenital immunodeficiency, or acquired immune deficiency, or patients
CC having had a bone marrow transplant or the elderly. This is the amino
CC acid sequence of a respiratory syncytial virus (RSV) antibody peptide
XX
SQ Sequence 7 AA;

Query Match 64.5%; Score 20; DB 6; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLKLS 6
|:|||
Db 2 TMKLS 6

RESULT 3
ADE35876
ID ADE35876 standard; peptide; 7 AA.
XX
AC ADE35876;
XX
DT 29-JAN-2004 (first entry)
XX
DE SYNAGIS antibody based light chain variable region 2, VL2, CDR #54.
XX
KW respiratory syncytial virus; RSV infection; high affinity antibody;
KW high avidity antibody; low antibody dose; more effective prophylaxis;
KW complementarity determining region; CDR; human.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN US2003091584-A1.
XX
PD 15-MAY-2003.
XX
PF 28-NOV-2001; 2001US-00996265.
XX
PR 28-NOV-2000; 2000US-00724396.
XX
PA (YOUN/) YOUNG J F.
PA (KOEN/) KOENIG S.
PA (JOHN/) JOHNSON L S.
XX
PI Young JF, Koenig S, Johnson LS;
XX
DR WPI; 2003-874589/81.
XX
PT Preventing, treating or ameliorating symptoms associated with respiratory
PT syncytial virus infection in mammal by administering antibodies or their
PT fragments that immunospecifically bind to RSV antigens.
XX
PS Claim 111; SEQ ID NO 164; 161pp; English.
XX
CC The invention relates to a method of preventing, treating or ameliorating
CC one or more symptoms associated with a respiratory syncytial virus (RSV)
CC infection in a mammal. A sustained release formulation comprising one or
CC more antibodies or their fragments that immunospecifically bind to one or
CC more RSV antigens and a pharmaceutical composition comprising one or more
CC antibodies or their fragments that immunospecifically bind to one or more
CC RSV antigens formulated for pulmonary delivery is useful for preventing,
CC treating or ameliorating one or more symptoms associated with a RSV
CC infection in a mammal by administering the formulation to the mammal, or
CC the composition to the lungs of the mammal. The antibodies bind to RSV
CC antigen with a high affinity and/or high avidity. The methods use lower
CC doses of antibodies which immunospecifically bind to RSV antigen, and
CC which provide a more effective prophylaxis. The present sequence
CC represents the amino acid sequence of a complementarity determining
CC region based on the SYNAGIS antibody.
XX
SQ Sequence 7 AA;

Query Match 64.5%; Score 20; DB 7; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLKLS 6
|:|||
Db 2 TMKLS 6

```
RESULT 4
AAW69269
ID   AAW69269 standard; peptide; 7 AA.
XX
AC   AAW69269;
XX
DT   29-OCT-1998  (first entry)
XX
DE   Haemagglutinin heavy chain (HA1) fragment.
XX
KW   Acryloylated peptide polymer; immune response; peptide epitope;
KW   synthetic vaccine; enzymatically cleavable site.
XX
OS   Influenza virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "linked to acryloylated peptide polymer"
XX
PN   WO9834968-A1.
XX
PD   13-AUG-1998.
XX
PF   10-FEB-1998;   98WO-AU000076.
XX
PR   11-FEB-1997;   97AU-00005071.
PR   03-OCT-1997;   97CA-02217321.
XX
PA   (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (UYME ) UNIV MELBOURNE.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PA   (CSLC-) CSL LTD.
XX
PI   Jackson DC,  Obrien-Simpson NM,  Brown LE,  Zeng W,  Ede NJ;
PI   Brandt ER,  Good MF;
XX
DR   WPI; 1998-447177/38.
XX
PT   Acryloylated peptide polymers - useful for synthetic vaccine technology,
PT   for raising an immune response to peptide epitope and as diagnostic tool.
XX
PS   Example 1; Page 20; 77pp; English.
XX
CC   This sequence represents a fragment of the heavy chain (HA1) of the
CC   haemagglutinin of influenza virus. This sequence was used to test the
CC   acryloylated peptide polymer of the invention. The peptide polymers are
CC   used to raise an immune response to a peptide epitope (such as this
CC   sequence), and also as diagnostic tools. Polymers (molecular wt. >600
CC   kDa.) can be prepared with virtually any number of the same or different
CC   epitopes by a method that allows purification of the individual
CC   determinants, avoids errors inherent in long sequential syntheses in
CC   which protected peptide fragments are not used, thus avoiding solubility
CC   and purification problems. Multiple copies of many different peptide
CC   epitopes may be incorporated into a single polymeric structure to allow
CC   utilisation of the range of T cell epitopes required for outbred
CC   populations in conjunction with epitopes representing different
CC   pathogenic serodemes, thus making them a significant advance in synthetic
CC   vaccine technology
XX
SQ   Sequence 7 AA;

  Query Match             61.3%;  Score 19;  DB 2;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLKLS 6
              ||||:
Db          1 TLKLA 5
```

```
RESULT 5
AAY42013
ID   AAY42013 standard; peptide; 7 AA.
XX
AC   AAY42013;
XX
DT   09-DEC-1999  (first entry)
XX
DE   Rheumatoid arthritis diagnostic protein isoform peptide #164.
XX
KW   Human; rheumatoid arthritis; RA; diagnosis; RPI; RADF; detection;
KW   rheumatoid arthritis diagnostic feature; ERPI; synovial fluid;
KW   rheumatoid arthritis diagnostic protein isoform; screening;
KW   expression reference protein isoform; prognosis.
XX
OS   Homo sapiens.
XX
PN   WO9947925-A2.
XX
PD   23-SEP-1999.
XX
PF   15-MAR-1999;   99WO-GB000763.
XX
PR   13-MAR-1998;   98GB-00005477.
XX
PA   (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI   Parekh RB,  Patel TP,  Townsend RR;
XX
DR   WPI; 1999-571871/48.
XX
PT   Diagnosis of human rheumatoid arthritis by two-dimensional
PT   electrophoresis.
XX
PS   Disclosure; Page 21; 157pp; English.
XX
CC   A method has been developed for the diagnosis of human rheumatoid
CC   arthritis (RA) using two-dimensional electrophoresis to generate a two-
CC   dimensional array of features. The method can be used for screening,
CC   diagnosis and prognosis of RA in a subject or for monitoring the effect
CC   of an anti-RA drug or therapy administered to a subject. The method
CC   comprises: (a) analysing a sample of serum or plasma and optionally
CC   synovial fluid by two-dimensional electrophoresis, to generate a two-
CC   dimensional array of features; (b) identifying at least one chosen
CC   feature whose relative abundance correlates with the presence or absence
CC   of RA; and (c) comparing the abundance of each chosen feature in the
CC   sample with the abundance of that chosen feature in serum or plasma from
CC   one or more persons without RA, where the relative abundance of the
CC   chosen feature or features in the sample indicates the presence or
CC   absence of RA in the subject. The method can also be used in clinical
CC   studies for testing drugs for therapy of RA, for purification of RA-
CC   diagnostic protein isoforms (RPIs), and for production of antibodies to
CC   RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify
CC   compounds that promote or inhibit their activity, which are then used as
CC   RA drugs. Nucleic acid encoding RADFs can be used in gene therapy
CC   protocols. AAY41844 to AAY42100 represent RPI peptides, AAY42101 to
CC   AAY42103 represent expression reference protein isoform peptides and
CC   AAZ25066 to AAZ25068 represent degenerate probes for RPIs, which are all
CC   used in the exemplification of the present invention
XX
SQ   Sequence 7 AA;

  Query Match             61.3%;  Score 19;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKLSR 7
              || :||
Db          2 TLMISR 7


RESULT 6
AAY41889
```

ID AAY41889 standard; peptide; 7 AA.
XX
AC AAY41889;
XX
DT 09-DEC-1999 (first entry)
XX
DE Rheumatoid arthritis diagnostic protein isoform peptide #40.
XX
KW Human; rheumatoid arthritis; RA; diagnosis; RPI; RADF; detection;
KW rheumatoid arthritis diagnostic feature; ERPI; synovial fluid;
KW rheumatoid arthritis diagnostic protein isoform; screening;
KW expression reference protein isoform; prognosis.
XX
OS Homo sapiens.
XX
PN WO9947925-A2.
XX
PD 23-SEP-1999.
XX
PF 15-MAR-1999; 99WO-GB000763.
XX
PR 13-MAR-1998; 98GB-00005477.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Parekh RB, Patel TP, Townsend RR;
XX
DR WPI; 1999-571871/48.
XX
PT Diagnosis of human rheumatoid arthritis by two-dimensional
PT electrophoresis.
XX
PS Disclosure; Page 18; 157pp; English.
XX
CC A method has been developed for the diagnosis of human rheumatoid
CC arthritis (RA) using two-dimensional electrophoresis to generate a two-
CC dimensional array of features. The method can be used for screening,
CC diagnosis and prognosis of RA in a subject or for monitoring the effect
CC of an anti-RA drug or therapy administered to a subject. The method
CC comprises: (a) analysing a sample of serum or plasma and optionally
CC synovial fluid by two-dimensional electrophoresis, to generate a two-
CC dimensional array of features; (b) identifying at least one chosen
CC feature whose relative abundance correlates with the presence or absence
CC of RA; and (c) comparing the abundance of each chosen feature in the
CC sample with the abundance of that chosen feature in serum or plasma from
CC one or more persons without RA, where the relative abundance of the
CC chosen feature or features in the sample indicates the presence or
CC absence of RA in the subject. The method can also be used in clinical
CC studies for testing drugs for therapy of RA, for purification of RA-
CC diagnostic protein isoforms (RPIs), and for production of antibodies to
CC RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify
CC compounds that promote or inhibit their activity, which are then used as
CC RA drugs. Nucleic acid encoding RADFs can be used in gene therapy
CC protocols. AAY41844 to AAY42100 represent RPI peptides, AAY42101 to
CC AAY42103 represent expression reference protein isoform peptides and
CC AAZ25066 to AAZ25068 represent degenerate probes for RPIs, which are all
CC used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 61.3%; Score 19; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLKLSR 7
| :|||
Db 2 TFELSR 7

RESULT 7
ABB55870
ID ABB55870 standard; peptide; 7 AA.
XX
AC ABB55870;
XX
DT 15-FEB-2002 (first entry)
XX
DE Vascular dementia-associated protein isoform (VPI) 70.
XX
KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW diagnosis; prognosis; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200169261-A2.
XX
PD 20-SEP-2001.
XX
PF 14-MAR-2001; 2001WO-GB001106.
XX
PR 15-MAR-2000; 2000GB-00006285.
PR 24-NOV-2000; 2000GB-00028734.
PR 28-NOV-2000; 2000US-00724391.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-557937/62.
XX
PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT determining stage of VD and monitoring the effect of VD therapy,
PT comprises analyzing body fluid by 2-dimensional electrophoresis for
PT features correlated with VD.
XX
PS Claim 6; Page 31; 151pp; English.
XX
CC The invention relates to screening, diagnosis or prognosis of Vascular
CC Dementia (VD) in a subject comprising analysing body fluid from the
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC features containing at least one chosen feature whose relative abundance
CC correlates with the presence, absence, stage or severity of VD or
CC predicts the onset or course of VD, especially detecting in a sample of
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC specification. Detecting VD-associated features and VPI is useful for the
CC screening, diagnosis or prognosis of VD, for determining the stage or
CC severity of VD, for identifying a subject at risk of VD or for monitoring
CC the effect of therapy administered to a subject having VD. Nucleic acids
CC encoding a VPI or inhibiting the function of a VPI are useful for the
CC treatment of VD and for gene therapy
XX
SQ Sequence 7 AA;

Query Match 61.3%; Score 19; DB 4; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLKLSR 7
| :|||
Db 2 TFELSR 7

RESULT 8
ABB56283
ID ABB56283 standard; peptide; 7 AA.
XX
AC ABB56283;
XX
DT 15-FEB-2002 (first entry)
XX
DE Vascular dementia-associated protein isoform (VPI) 483.
XX
KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW diagnosis; prognosis; gene therapy.
XX

OS Homo sapiens.
XX
PN WO200169261-A2.
XX
PD 20-SEP-2001.
XX
PF 14-MAR-2001; 2001WO-GB001106.
XX
PR 15-MAR-2000; 2000GB-00006285.
PR 24-NOV-2000; 2000GB-00028734.
PR 28-NOV-2000; 2000US-00724391.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-557937/62.
XX
PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT determining stage of VD and monitoring the effect of VD therapy,
PT comprises analyzing body fluid by 2-dimensional electrophoresis for
PT features correlated with VD.
XX
PS Claim 6; Page 40; 151pp; English.
XX
CC The invention relates to screening, diagnosis or prognosis of Vascular
CC Dementia (VD) in a subject comprising analysing body fluid from the
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC features containing at least one chosen feature whose relative abundance
CC correlates with the presence, absence, stage or severity of VD or
CC predicts the onset or course of VD, especially detecting in a sample of
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC specification. Detecting VD-associated features and VPI is useful for the
CC screening, diagnosis or prognosis of VD, for determining the stage or
CC severity of VD, for identifying a subject at risk of VD or for monitoring
CC the effect of therapy administered to a subject having VD. Nucleic acids
CC encoding a VPI or inhibiting the function of a VPI are useful for the
CC treatment of VD and for gene therapy
XX
SQ Sequence 7 AA;

Query Match 61.3%; Score 19; DB 4; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLKLSR 7
| :|||
Db 2 TFELSR 7

RESULT 9
ABB55981
ID ABB55981 standard; peptide; 7 AA.
XX
AC ABB55981;
XX
DT 15-FEB-2002 (first entry)
XX
DE Vascular dementia-associated protein isoform (VPI) 181.
XX
KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW diagnosis; prognosis; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200169261-A2.
XX
PD 20-SEP-2001.
XX
PF 14-MAR-2001; 2001WO-GB001106.
XX
PR 15-MAR-2000; 2000GB-00006285.
PR 24-NOV-2000; 2000GB-00028734.
PR 28-NOV-2000; 2000US-00724391.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-557937/62.
XX
PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT determining stage of VD and monitoring the effect of VD therapy,
PT comprises analyzing body fluid by 2-dimensional electrophoresis for
PT features correlated with VD.
XX
PS Claim 6; Page 33; 151pp; English.
XX
CC The invention relates to screening, diagnosis or prognosis of Vascular
CC Dementia (VD) in a subject comprising analysing body fluid from the
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC features containing at least one chosen feature whose relative abundance
CC correlates with the presence, absence, stage or severity of VD or
CC predicts the onset or course of VD, especially detecting in a sample of
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC specification. Detecting VD-associated features and VPI is useful for the
CC screening, diagnosis or prognosis of VD, for determining the stage or
CC severity of VD, for identifying a subject at risk of VD or for monitoring
CC the effect of therapy administered to a subject having VD. Nucleic acids
CC encoding a VPI or inhibiting the function of a VPI are useful for the
CC treatment of VD and for gene therapy
XX
SQ Sequence 7 AA;

Query Match 61.3%; Score 19; DB 4; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLKLSR 7
| :|||
Db 2 TFELSR 7

RESULT 10
AAU28602
ID AAU28602 standard; peptide; 7 AA.
XX
AC AAU28602;
XX
DT 03-JAN-2002 (first entry)
XX
DE DPI tryptic digest peptide #199.
XX
KW Human; depression associated protein isoform; tryptic digest peptide;
KW DPI; cerebrospinal fluid; CSF; BAD; bipolar affective disorder;
KW neuropsychiatric disorder; bipolar mood disorder; neuroleptic;
KW maniac-depressive illness; schizoaffective disorder.
XX
OS Homo sapiens.
XX
PN WO200162787-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000786.
XX
PR 24-FEB-2000; 2000GB-00004412.
PR 08-DEC-2000; 2000GB-00030050.
PR 12-DEC-2000; 2000US-0254830P.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C, Terrett JA, Tyson KL;
XX

DR WPI; 2001-570626/64.
XX
PT Novel nucleic acid encoding a protein associated with bipolar affective
PT disorder, which is used for diagnosis, prophylaxis and therapy of
PT neuropsychiatric disorders, such as bipolar affective disorder.
XX
PS Disclosure; Page 34; 153pp; English.
XX
CC The present invention relates to the identification of depression
CC associated protein isoforms (DPIs), particularly the tryptic digest
CC peptides of these proteins. Some of the DPIs (AAU28404-AAU28625)
CC described are decreased in the cerebrospinal fluid (CSF) of BAD (bipolar
CC affective disorder) subjects, whilst other DPIs (AAU28626-AAU28887) are
CC increased in BAD subjects. Also described are peptide sequences
CC identified from DPI-45 and DPI-213 and the nucleic acid sequence they are
CC encoded by. The sequences of the invention are useful for clinical
CC screening, diagnosis, prognosis, therapy and prophylaxis of
CC neuropsychiatric disorders e.g. BAD (also known as bipolar mood disorder,
CC BP), maniac-depressive illnesses, attention deficit disorders,
CC schizoaffective disorders, and unipolar affective disorders. The present
CC sequence represents one of the DPI tryptic digest peptides of the present
CC invention
XX
SQ Sequence 7 AA;

Query Match 61.3%; Score 19; DB 4; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLKLSR 7
| :|||
Db 2 TFELSR 7

RESULT 11
AAU24969
ID AAU24969 standard; peptide; 7 AA.
XX
AC AAU24969;
XX
DT 18-DEC-2001 (first entry)
XX
DE Schizophrenia-Associated Protein Isoform (SPI) peptide #198.
XX
KW Schizophrenia-associated protein isoform; SPI; SPI-206; SPI-238; SPI-240;
KW neuroleptic; gene therapy; cerebrospinal fluid; serum; plasma.
XX
OS Homo sapiens.
XX
PN WO200162785-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000792.
XX
PR 24-FEB-2000; 2000GB-00004415.
PR 28-DEC-2000; 2000US-00750395.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C, Terrett JA, Tyson KL;
XX
DR WPI; 2001-570624/64.
XX
PT New schizophrenia associated protein isoforms and encoding nucleic acid
PT molecules, useful for treatment, diagnosis and prognosis of schizophrenia
PT and screening for potential drugs for treatment and new drug targets.
XX
PS Disclosure; Page 32; 148pp; English.
XX
CC The sequence represents a schizophrenia-associated protein isoform (SPI).
CC These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable
CC in cerebrospinal fluid, serum or plasma and are useful markers of
CC schizophrenia. The sequences can be used for treatment and diagnosis of
CC schizophrenia, screening, prognosis, monitoring the results of therapy,
CC identifying patients most likely to respond to a particular therapy and
CC identification of new targets for drug treatment. SPI DNA is useful as a
CC nucleic acid probe to detect the presence of nucleic acids or SPIs
XX
SQ Sequence 7 AA;

Query Match 61.3%; Score 19; DB 4; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLKLSR 7
| :|||
Db 2 TFELSR 7

RESULT 12
AAU26249
ID AAU26249 standard; peptide; 7 AA.
XX
AC AAU26249;
XX
DT 18-DEC-2001 (first entry)
XX
DE Depression-Associated Protein isoform DPI-208.
XX
KW Human; Bipolar Affective Disorder; BAD; Depression-Associated feature;
KW DF; Depression-Associated protein isoform; DPI; Cerebro-spinal fluid;
KW CSF; antidepressant; antimanic; nootropic; tranquiliser; neuroleptic;
KW attention deficient disorder; schizoaffective disorder;
KW unipolar affective disorder.
XX
OS Homo sapiens.
XX
PN WO200163294-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000791.
XX
PR 24-FEB-2000; 2000GB-00004412.
PR 08-DEC-2000; 2000GB-00030050.
PR 12-DEC-2000; 2000US-0254830P.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-582081/65.
XX
PT Preparation for diagnosing or treating bipolar affected disorder (BAD) or
PT unipolar depression, or for screening for modulators, comprises a BAD-
PT associated protein isoform.
XX
PS Claim 8; Page 34; 163pp; English.
XX
CC The invention relates to a preparation comprising an isolated Bipolar
CC Affected Disorder (BAD)-Associated Protein Isoform (DPIs). The DPI's are
CC used to screen, diagnose or prognose of BAD or unipolar depression,
CC determine the stage or severity of BAD or unipolar depression, identify a
CC subject at risk of developing BAD or unipolar depression, or monitor the
CC effect of therapy in a subject. They are also used to screen for or
CC identify agents that interact with a DPI. These agents, antibodies
CC against the DPIs, and nucleic acids encoding the DPIs are used to treat
CC or prevent BAD or unipolar depression. Diseases that can be treated are
CC attention deficient disorder, a schizoaffective disorder, a bipolar or a
CC unipolar affective disorder. The DPIs are used in proteomics. The
CC proteomic approach of using DPIs for screening, diagnosis or prognosis of
CC BAD or unipolar depression overcomes the problems of using gene
CC expression analysis, such as not being able to obtain central nervous
CC system (CNS) tissue from a living patient under normal circumstances. The
CC present sequence is a DIP decreased in the CSF (cerebro-spinal fluid) of

```
CC   subjects having BAD
XX
SQ   Sequence 7 AA;

  Query Match             61.3%;  Score 19;  DB 4;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKLSR 7
              | :|||
Db          2 TFELSR 7


RESULT 13
AAU15313
ID   AAU15313 standard; peptide; 7 AA.
XX
AC   AAU15313;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Schizophrenia-associated isoform peptide #198.
XX
KW   Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder;
KW   neurological disorder; neuropathy.
XX
OS   Homo sapiens.
XX
PN   WO200163293-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000783.
XX
PR   24-FEB-2000; 2000GB-00004415.
PR   28-DEC-2000; 2000US-00750395.
XX
PA   (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI   Herath HMAC,  Parekh RB,  Rohlff C;
XX
DR   WPI; 2001-502868/55.
XX
PT   Diagnosing and monitoring Schizophrenia by detecting the presence of
PT   Schizophrenia Associated Features and Schizophrenia Associated Protein
PT   Isoforms in samples of cerebrospinal fluid.
XX
PS   Claim 6; Page 32; 160pp; English.
XX
CC   The invention relates to methods and compositions for screening,
CC   diagnosis and prognosis of Schizophrenia. The method involves detecting
CC   the presence of Schizophrenia (SCH) Associated Features (SFs) and SCH
CC   Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis,
CC   immunoassay or hybridisation assay, for diagnosing and monitoring SCH,
CC   studying the effectiveness of treatments and for identifying potential
CC   therapeutic agents. The method is used for (1) screening or diagnosis of
CC   SCH and the relative abundance of at least 1 chosen feature correlates
CC   with the presence or absence of SCH; and (2) monitoring the effect of
CC   therapy administered to a subject with SCH and the relative abundance of
CC   at least 1 chosen feature which correlates with the severity of SCH. The
CC   expression and activity of the SFs, SPIs and related molecules (e.g.
CC   secondary messengers) are studied to diagnose SCH, monitor the progress
CC   of the disorder and the effectiveness of treatment and as targets to
CC   identify and produce potential therapeutic agents for the treatment of
CC   SCH. The paucity of detectable neuralgic defects distinguishes
CC   neuropsychiatric disorders such as SCH from neurological disorders, where
CC   manifestations of anatomical and biochemical changes have been identified
CC   in many cases. Consequently the identification and characterisation of
CC   cellular and/or molecular causative defects and neuropathies are
CC   necessary for improved treatment of neuropsychiatric disorders. AAU15114-
CC   AAU15762 represent the amino acid sequences of schizophrenia-associated
CC   isoforms used in the method of the invention
XX
SQ   Sequence 7 AA;

  Query Match             61.3%;  Score 19;  DB 4;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKLSR 7
              | :|||
Db          2 TFELSR 7


RESULT 14
ABB52355
ID   ABB52355 standard; peptide; 7 AA.
XX
AC   ABB52355;
XX
DT   08-FEB-2002  (first entry)
XX
DE   Human API-125 tryptic digest peptide #8.
XX
KW   Human; neuroprotective; nootropic; gene therapy; vaccine;
KW   Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;
KW   Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;
KW   Expression Reference Protein Isoform; ERPI; proteolysis.
XX
OS   Homo sapiens.
XX
PN   WO200175454-A2.
XX
PD   11-OCT-2001.
XX
PF   03-APR-2001; 2001WO-US010908.
XX
PR   03-APR-2000; 2000US-0194504P.
PR   28-NOV-2000; 2000US-0253647P.
XX
PA   (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
PA   (PFIZ ) PFIZER INC.
XX
PI   Durham KL,  Friedman DL,  Herath HMAC,  Kimmel LH,  Parekh RB;
PI   Potter DM,  Rohlff C,  Silber BM,  Stiger TR,  Sunderland PT;
PI   Townsend RR,  White F,  Williams SA;
XX
DR   WPI; 2001-639384/73.
XX
PT   Screening for Alzheimer's disease in a mammal, by making two-dimensional
PT   array of a feature whose relative abundance correlates with disease, and
PT   comparing with abundance of the feature in samples of healthy persons.
XX
PS   Example; Page 34; 162pp; English.
XX
CC   The invention relates to methods for the screening, diagnosis and
CC   prognosis of Alzheimer's disease. The methods involve the detection of
CC   Alzheimer's Disease-Associated Features (AFs) and Alzheimer's Disease-
CC   Associated Protein Isoforms (APIs) in cerebrospinal fluid, serum or
CC   plasma. The abundance of the AFs and APIs is then normalised to an
CC   Expression Reference Protein Isoform (ERPI) in order to determine whether
CC   a patient is suffering from, or has a predisposition to, Alzheimer's
CC   Disease. The relative abundance of the AFs and APIs correlates with the
CC   severity of Alzheimer's Disease. The present sequence is a peptide
CC   produced from an API by proteolysis
XX
SQ   Sequence 7 AA;

  Query Match             61.3%;  Score 19;  DB 4;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKLSR 7
              | :|||
Db          2 TFELSR 7
```

RESULT 15
ABP66489
ID ABP66489 standard; peptide; 7 AA.
XX
AC ABP66489;
XX
DT 04-DEC-2002 (first entry)
XX
DE Human RSV antibody VL CDR2 fragment.
XX
KW Human; variable heavy domain; variable light domain; CDR; VH; VL; RSV;
KW complementarity determining region; respiratory syncytial virus;
KW virucide; pulmonary; antiinflammatory; cardiant; anti-HIV; vaccine;
KW immunostimulant; gene therapy; cystic fibrosis; bone marrow transplant;
KW bronchopulmonary dysplasia; congenital heart disease;
KW congenital immunodeficiency; acquired immunodeficiency.
XX
OS Homo sapiens.
XX
PN WO200243660-A2.
XX
PD 06-JUN-2002.
XX
PF 28-NOV-2001; 2001WO-US044807.
XX
PR 28-NOV-2000; 2000US-00724396.
PR 28-NOV-2000; 2000US-00724531.
XX
PA (MEDI-) MEDIUMMUNE INC.
XX
PI Young JF, Koenig S, Johnson LS;
XX
DR WPI; 2002-706803/76.
XX
PT Antibody for treating respiratory syncytial virus (RSV) infection,
PT comprises a variable heavy/light domain or complementarity determining
PT regions 1 - 3 of variable light/heavy chains, that immunospecifically
PT binds to RSV antigen.
XX
PS Claim 8; Page 55; 298pp; English.
XX
CC The invention relates to a novel antibody comprising a variable heavy
CC (VH) domain, variable light (VL) domain, VH complementarity determining
CC region (CDR)-1, VH CDR2, VH CDR3, VL CDR1, VL CDR2 or VL CDR3, where the
CC antibody immunospecifically binds to a respiratory syncytial virus (RSV)
CC antigen, and where the antibody is not SYNAGIS (RTM). The antibody of the
CC invention has virucide, pulmonary, antiinflammatory, cardiant, anti-HIV,
CC and immunostimulant activity. The polynucleotides of the invention may
CC have a use in a vaccine, and in gene therapy. The antibody is useful for
CC treating or ameliorating a RSV infection in a human. The antibody is also
CC useful for preventing, treating or ameliorating one or more symptoms
CC associated with RSV infection in a mammal, e.g. cystic fibrosis,
CC bronchopulmonary dysplasia, congenital heart disease, congenital
CC immunodeficiency or acquired immunodeficiency, or after a bone marrow
CC transplant. The sequence represents a complementary determining region
CC peptide from a human RSV antibody of the invention
XX
SQ Sequence 7 AA;

Query Match 61.3%; Score 19; DB 5; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 TLKLS 6
              ||||:
Db          2 TLKLA 6
```

Search completed: June 2, 2004, 18:57:51
Job time : 50 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2004, 18:59:52 ; Search time 37.5 Seconds
(without alignments)
52.517 Million cell updates/sec

Title: US-09-712-819D-13
Perfect score: 31
Sequence: 1 LTLKLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 49349

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	64.5	7	9	US-09-996-288-164	Sequence 164, App
2	20	64.5	7	10	US-09-996-265-164	Sequence 164, App
3	20	64.5	7	15	US-10-461-863-164	Sequence 164, App
4	19	61.3	7	9	US-09-791-378-198	Sequence 198, App
5	19	61.3	7	9	US-09-998-909-7	Sequence 7, Appli
6	19	61.3	7	9	US-09-826-290-370	Sequence 370, App
7	19	61.3	7	9	US-09-996-288-135	Sequence 135, App
8	19	61.3	7	10	US-09-791-393-145	Sequence 145, App
9	19	61.3	7	10	US-09-791-389-145	Sequence 145, App
10	19	61.3	7	10	US-09-996-265-135	Sequence 135, App
11	19	61.3	7	12	US-10-601-100-105	Sequence 105, App
12	19	61.3	7	13	US-10-044-034-13	Sequence 13, Appl
13	19	61.3	7	15	US-10-461-863-135	Sequence 135, App
14	19	61.3	7	16	US-10-264-309-460	Sequence 460, App
15	18	58.1	6	9	US-09-897-107-24	Sequence 24, Appl
16	18	58.1	7	9	US-09-989-789-396	Sequence 396, App
17	18	58.1	7	9	US-09-989-789-1971	Sequence 1971, Ap
18	18	58.1	7	9	US-09-989-789-3503	Sequence 3503, Ap
19	18	58.1	7	9	US-09-989-789-3542	Sequence 3542, Ap
20	18	58.1	7	9	US-09-989-789-3544	Sequence 3544, Ap
21	18	58.1	7	9	US-09-796-848A-20	Sequence 20, Appl
22	18	58.1	7	9	US-09-844-508-43	Sequence 43, Appl
23	18	58.1	7	9	US-09-996-288-35	Sequence 35, Appl
24	18	58.1	7	9	US-09-996-288-96	Sequence 96, Appl
25	18	58.1	7	9	US-09-996-288-131	Sequence 131, App
26	18	58.1	7	10	US-09-990-186-396	Sequence 396, App
27	18	58.1	7	10	US-09-990-186-1971	Sequence 1971, Ap
28	18	58.1	7	10	US-09-990-186-3503	Sequence 3503, Ap
29	18	58.1	7	10	US-09-990-186-3542	Sequence 3542, Ap
30	18	58.1	7	10	US-09-990-186-3544	Sequence 3544, Ap
31	18	58.1	7	10	US-09-996-265-35	Sequence 35, Appl
32	18	58.1	7	10	US-09-996-265-96	Sequence 96, Appl
33	18	58.1	7	10	US-09-996-265-131	Sequence 131, App
34	18	58.1	7	10	US-09-989-994-396	Sequence 396, App
35	18	58.1	7	10	US-09-989-994-1971	Sequence 1971, Ap
36	18	58.1	7	10	US-09-989-994-3503	Sequence 3503, Ap
37	18	58.1	7	10	US-09-989-994-3542	Sequence 3542, Ap
38	18	58.1	7	10	US-09-989-994-3544	Sequence 3544, Ap
39	18	58.1	7	14	US-10-084-826-43	Sequence 43, Appl
40	18	58.1	7	14	US-10-234-026-6	Sequence 6, Appli
41	18	58.1	7	14	US-10-020-354-35	Sequence 35, Appl
42	18	58.1	7	15	US-10-461-863-35	Sequence 35, Appl
43	18	58.1	7	15	US-10-461-863-96	Sequence 96, Appl
44	18	58.1	7	15	US-10-461-863-131	Sequence 131, App
45	17	54.8	7	9	US-09-996-288-121	Sequence 121, App

ALIGNMENTS

RESULT 1
US-09-996-288-164
; Sequence 164, Application US/09996288
; Patent No. US20020177126A1
; GENERAL INFORMATION:
; APPLICANT: Young, James
; APPLICANT: Scott, Koenig
; APPLICANT: Leslie, Johnson
; TITLE OF INVENTION: Methods of Adminstering/Dosing Anti-RSV Antibodies for Prophyl
; TITLE OF INVENTION: and Treatment
; FILE REFERENCE: 10271-047-999
; CURRENT APPLICATION NUMBER: US/09/996,288
; CURRENT FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-288-164

Query Match 64.5%; Score 20; DB 9; Length 7;
Best Local Similarity 80.0%; Pred. No. 1e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLKLS 6
|:|||
Db 2 TMKLS 6

RESULT 2
US-09-996-265-164
; Sequence 164, Application US/09996265
; Publication No. US20030091584A1
; GENERAL INFORMATION:
; APPLICANT: Young, James
; APPLICANT: Scott, Koenig
; APPLICANT: Leslie, Johnson

; TITLE OF INVENTION: Methods of Adminstering/Dosing Anti-RSV Antibodies for Prophylaxi
; TITLE OF INVENTION: and Treatment
; FILE REFERENCE: 10271-048-999
; CURRENT APPLICATION NUMBER: US/09/996,265
; CURRENT FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-265-164

Query Match 64.5%; Score 20; DB 10; Length 7;
Best Local Similarity 80.0%; Pred. No. 1e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLKLS 6
|:|||
Db 2 TMKLS 6

RESULT 3
US-10-461-863-164
; Sequence 164, Application US/10461863
; Publication No. US20040018200A1
; GENERAL INFORMATION:
; APPLICANT: Oliver, Cynthia
; APPLICANT: Allan, Christian
; APPLICANT: Chang, Stephen
; TITLE OF INVENTION: STABILIZED ANTI-RESPIRATORY SYNCYTIAL VIRUS (RSV) ANTIBODY FORMUL
; FILE REFERENCE: 10271-071-999
; CURRENT APPLICATION NUMBER: US/10/461,863
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: 60/388,920
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 209
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-461-863-164

Query Match 64.5%; Score 20; DB 15; Length 7;
Best Local Similarity 80.0%; Pred. No. 1e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLKLS 6
|:|||
Db 2 TMKLS 6

RESULT 4
US-09-791-378-198
; Sequence 198, Application US/09791378
; Patent No. US20020142303A1
; GENERAL INFORMATION:
; APPLICANT: Parekh, Rajesh
; TITLE OF INVENTION: PROTEINS, GENES AND THEIR USE FOR DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: SCHIZOPHRENIA
; FILE REFERENCE: 9195-061-999
; CURRENT APPLICATION NUMBER: US/09/791,378
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/750,395
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 677
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 198
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-378-198

Query Match 61.3%; Score 19; DB 9; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLKLSR 7
| :|||
Db 2 TFELSR 7

RESULT 5
US-09-998-909-7
; Sequence 7, Application US/09998909
; Patent No. US20020164664A1
; GENERAL INFORMATION:
; APPLICANT: Hlavaty, John
; APPLICANT: Briggman, Joseph
; TITLE OF INVENTION: Detection and Treatment of Prostate Cancer
; FILE REFERENCE: MTP-027
; CURRENT APPLICATION NUMBER: US/09/998,909
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 60/250,284
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-998-909-7

Query Match 61.3%; Score 19; DB 9; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLKLSR 7
| :|||
Db 2 TFELSR 7

RESULT 6
US-09-826-290-370
; Sequence 370, Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
; APPLICANT: Durham, L.Kathryn
; APPLICANT: Friedman, David L.
; APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
; APPLICANT: Kimmel, Lida H.
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Potter, David M.
; APPLICANT: Rohlff, Christian
; APPLICANT: Silber, B. Michael
; APPLICANT: Stiger, Thomas R.
; APPLICANT: Sunderland, P. Trey
; APPLICANT: Townsend, Robert Reid
; APPLICANT: White, Frost
; APPLICANT: Williams, Stephen A.
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
; TITLE OF INVENTION: Uses Therefor, Including Diagnosis and Treatment of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 2572-1-001 N2
; CURRENT APPLICATION NUMBER: US/09/826,290
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/194,504
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/253,647
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 492
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 370
; LENGTH: 7
; TYPE: PRT

```
;   ORGANISM: homo sapien
US-09-826-290-370

  Query Match             61.3%;  Score 19;  DB 9;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKLSR 7
              | :|||
Db          2 TFELSR 7


RESULT 7
US-09-996-288-135
; Sequence 135, Application US/09996288
; Patent No. US20020177126A1
; GENERAL INFORMATION:
;  APPLICANT: Young, James
;  APPLICANT:  Scott, Koenig
;  APPLICANT:  Leslie, Johnson
;  TITLE OF INVENTION: Methods of Adminstering/Dosing Anti-RSV Antibodies for Prophylaxi
;  TITLE OF INVENTION:  and Treatment
;  FILE REFERENCE: 10271-047-999
;  CURRENT APPLICATION NUMBER: US/09/996,288
;  CURRENT FILING DATE:  2001-11-28
;  NUMBER OF SEQ ID NOS: 259
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 135
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-996-288-135

  Query Match             61.3%;  Score 19;  DB 9;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLKLS 6
              ||||:
Db          2 TLKLA 6


RESULT 8
US-09-791-393-145
; Sequence 145, Application US/09791393
; Publication No. US20030032200A1
; GENERAL INFORMATION:
;  APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT: Parekh, Rajesh Bhikhu
;  APPLICANT: Rohlff, Christian
;  TITLE OF INVENTION: Proteins, Genes and Their Use for
;  TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
;  TITLE OF INVENTION: and Unipolar Depression
;  FILE REFERENCE: 2543-1-001 N1
;  CURRENT APPLICATION NUMBER: US/09/791,393
;  CURRENT FILING DATE: 2002-01-02
;  EARLIER APPLICATION NUMBER: GB 0004412.3
;  EARLIER FILING DATE: 2000-02-24
;  EARLIER APPLICATION NUMBER: GB 0030050.9
;  EARLIER FILING DATE: 2000-12-08
;  EARLIER APPLICATION NUMBER: US 60/254,830
;  EARLIER FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 145
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-791-393-145

  Query Match             61.3%;  Score 19;  DB 10;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKLSR 7
              | :|||
Db          2 TFELSR 7


RESULT 9
US-09-791-389-145
; Sequence 145, Application US/09791389
; Publication No. US20030032773A1
; GENERAL INFORMATION:
;  APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT:  Parekh, Rajesh Bhikhu
;  APPLICANT:  Rohlff, Christian
;  APPLICANT:  Terrett, Jonathan Alexander
;  APPLICANT:  Tyson, Kerry Louise
;  TITLE OF INVENTION: Proteins, Genes and Their Use for
;  TITLE OF INVENTION:  Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
;  TITLE OF INVENTION:  and Unipolar Depression
;  FILE REFERENCE: 2543-1-001 N2
;  CURRENT APPLICATION NUMBER: US/09/791,389
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: GB 0004412.3
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: GB 0030050.9
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: US 60/254,830
;  PRIOR FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 145
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-791-389-145

  Query Match             61.3%;  Score 19;  DB 10;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKLSR 7
              | :|||
Db          2 TFELSR 7


RESULT 10
US-09-996-265-135
; Sequence 135, Application US/09996265
; Publication No. US20030091584A1
; GENERAL INFORMATION:
;  APPLICANT: Young, James
;  APPLICANT:  Scott, Koenig
;  APPLICANT:  Leslie, Johnson
;  TITLE OF INVENTION: Methods of Adminstering/Dosing Anti-RSV Antibodies for Prophyl
;  TITLE OF INVENTION:  and Treatment
;  FILE REFERENCE: 10271-048-999
;  CURRENT APPLICATION NUMBER: US/09/996,265
;  CURRENT FILING DATE:  2001-11-28
;  NUMBER OF SEQ ID NOS: 259
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 135
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-996-265-135

  Query Match             61.3%;  Score 19;  DB 10;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLKLS 6
              ||||:
Db          2 TLKLA 6
```

```
RESULT 11
US-10-601-100-105
; Sequence 105, Application US/10601100
; Publication No. US20040072261A1
; GENERAL INFORMATION:
;  APPLICANT: INNOGENETICS N.V.
;  TITLE OF INVENTION: Method for the Diagnosis and Differential Diagnosis of
;  TITLE OF INVENTION:  Neurological Diseases
;  FILE REFERENCE: 11362.0038.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/601,100
;  CURRENT FILING DATE:  2003-06-20
;  PRIOR APPLICATION NUMBER: EP 02447121.1
;  PRIOR FILING DATE: 2002-06-21
;  PRIOR APPLICATION NUMBER: US 60/396,437
;  PRIOR FILING DATE: 2002-07-17
;  NUMBER OF SEQ ID NOS: 113
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 105
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-10-601-100-105

  Query Match             61.3%;  Score 19;  DB 12;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKLSR 7
              | :|||
Db          2 TFELSR 7


RESULT 12
US-10-044-034-13
; Sequence 13, Application US/10044034
; Publication No. US20020169264A1
; GENERAL INFORMATION:
;  APPLICANT: JACKSON, DAVID C.
;  APPLICANT:  O'BRIEN-SIMPSON, NEIL M.
;  APPLICANT:  BROWN, LORENA E.
;  APPLICANT:  EDE, NICHOLAS J.
;  APPLICANT:  BRANDT, EVELYN R.
;  APPLICANT:  GOOD, MICHAEL F.
;  TITLE OF INVENTION: POLYMERS INCORPORTING PEPTIDES
;  FILE REFERENCE: FBRC:006
;  CURRENT APPLICATION NUMBER: US/10/044,034
;  CURRENT FILING DATE:  2002-01-11
;  PRIOR APPLICATION NUMBER: P05071
;  PRIOR FILING DATE: 1997-02-11
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Peptides
US-10-044-034-13

  Query Match             61.3%;  Score 19;  DB 13;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLKLS 6
              ||||:
Db          1 TLKLA 5


RESULT 13
US-10-461-863-135
; Sequence 135, Application US/10461863
; Publication No. US20040018200A1
; GENERAL INFORMATION:
;  APPLICANT: Oliver, Cynthia
;  APPLICANT:  Allan, Christian
;  APPLICANT:  Chang, Stephen
;  TITLE OF INVENTION: STABILIZED ANTI-RESPIRATORY SYNCYTIAL VIRUS (RSV) ANTIBODY FORM
;  FILE REFERENCE: 10271-071-999
;  CURRENT APPLICATION NUMBER: US/10/461,863
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: 60/388,920
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 209
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 135
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-461-863-135

  Query Match             61.3%;  Score 19;  DB 15;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLKLS 6
              ||||:
Db          2 TLKLA 6


RESULT 14
US-10-264-309-460
; Sequence 460, Application US/10264309
; Publication No. US20040022794A1
; GENERAL INFORMATION:
;  APPLICANT: DURHAM, L. KATHRYN
;  APPLICANT:  FRIEDMAN, DAVID L.
;  APPLICANT:  HERATH, HERATH
;  APPLICANT:  KIMMEL, LIDA H.
;  APPLICANT:  PAREKH, RAJESH B.
;  APPLICANT:  POTTER, DAVID M.
;  APPLICANT:  ROHLFF, CHRISTIAN
;  APPLICANT:  SILBER, B. MICHAEL
;  APPLICANT:  SNYDER, PETER J.
;  APPLICANT:  SOARES, HOLLY D.
;  APPLICANT:  STIGER, THOMAS R.
;  APPLICANT:  SUNDERLAND, P. TREY
;  APPLICANT:  TOWNSEND, ROBERT R.
;  APPLICANT:  WHITE, W. FROST
;  APPLICANT:  WILLIAMS, STEPHEN A.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES, POLYPEPTIDES AND USES THEREFOR,
;  TITLE OF INVENTION:  INCLUDING DIAGNOSIS AND TREATMENT OF ALZHEIMER'S DISEASE
;  FILE REFERENCE: POA-002.01
;  CURRENT APPLICATION NUMBER: US/10/264,309
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: 60/326,708
;  PRIOR FILING DATE: 2001-10-03
;  NUMBER OF SEQ ID NOS: 491
;  SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 460
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-264-309-460

  Query Match             61.3%;  Score 19;  DB 16;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLKLSR 7
              | :|||
Db          2 TFELSR 7
```

```
RESULT 15
US-09-897-107-24
; Sequence 24, Application US/09897107
; Patent No. US20020137094A1
; GENERAL INFORMATION:
;  APPLICANT: YAMAGISHI, Akihiko
;  TITLE OF INVENTION: METHOD FOR IMPROVING THERMOSTABILITY OF PROTEINS, PROTEINS HAVING
;  TITLE OF INVENTION:  THERMOSTABILITY IMPROVED BY THE METHOD AND NUCLEIC ACIDS ENCODIN
;  FILE REFERENCE: 210383US0
;  CURRENT APPLICATION NUMBER: US/09/897,107
;  CURRENT FILING DATE:  2001-07-03
;  PRIOR APPLICATION NUMBER: JP2000-201920
;  PRIOR FILING DATE: 2000-07-04
;  PRIOR APPLICATION NUMBER: JP2001-164332
;  PRIOR FILING DATE: 2001-05-31
;  NUMBER OF SEQ ID NOS: 104
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-897-107-24

  Query Match             58.1%;  Score 18;  DB 9;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 1e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LTLKLS 6
              : ||||
Db          1 MMLKLS 6


Search completed: June  2, 2004, 19:08:55
Job time : 38.5 secs
```